Analysis and characterisation of human chorionic gonadotropin glycoforms in pregnancy and trophoblastic disorders

A thesis submitted to Middlesex University in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Marcela B Gondek

Department of Natural Sciences,
Faculty of Science and Technology,
Middlesex University, London

April 2022

Abstract

Human chorionic gonadotropin (hCG) is a heterogeneous glycoprotein hormone with a varying degree of carbohydrate moieties. Its glycosylation arrangements are increasingly recognized as a common and important element of disease pathophysiology and are associated with many disorders including gestational trophoblastic diseases (GTDs). This study aimed to optimise methodologies to permit the characterisation of hCG N-linked glycans from urine samples collected throughout normal pregnancy and GTD using matrix-assisted laser desorption ionisation time of flight mass spectrometry (MALDI-TOF MS). hCG isolated from pooled pregnancy urine was used in this study. All the stages in pregnancy urine preparation were optimised; including conditions for hCG immunopurification, deglycosylation, solid-phase extraction of resulting glycan:protein mixture and application of N-glycans for MALDI-TOF MS analysis. GlycoQuest software was used to characterise specific N-glycans configurations from the resulting MALDI-TOF MS spectra. This methodology was then applied to urine samples collected throughout normal singleton pregnancy and urine collected from patients with choriocarcinoma. The relative abundance of glycans of different molecular weights and specific types (i.e. fucosylated, sialylated, bisected and sulphated) at each stage of normal pregnancy and in GTD were compared. Each stage of optimisation increased the number of N-glycans detected such that we were ultimately able to detect 50 different glycans in normal pregnancy urine. In these samples, advancing gestation was associated with an increase in the proportion of branched N-glycans and multi-fucosylated N-glycans. Also, a significant increase in the proportion of high molecular weight glycans (>2100 Da) between choriocarcinoma and first-trimester normal pregnancy was observed. Further striking differences in the repertoire of glycan expression were also seen in choriocarcinoma urine compared with first-trimester pregnancy urine. The proportion of multi-fucosylated and tri-and tetra antennal, glycans was increased 3 and 2 fold respectively. In addition, 14 unique N-linked glycan structures were identified in choriocarcinoma samples. These included hyperfucosylated (7 fucose groups) and hypersialylated (4 sialic acid groups) glycans. A feature of this unique set of glycans was that they contained a combination of multiple branching, fucosylation, sialylation, sulphation and glycans with Lewis[X] terminal epitopes. In summary, we have successfully developed a methodology for the detection of a diverse range of N-linked glycans from hCG. These results suggest that this approach can be successfully used for the detection of novel glyco-biomarkers for the early detection of choriocarcinoma and may be applied to other GTDs associated with a dysregulation of hCG expression.

Attribution

MIDDLESEX UNIVERSITY FORM
Candidate's Declaration Form

Name of candidate: Marcela B Gondek
Thesis Title: Analysis and characterisation of human chorionic gonadotropin glycoforms in pregnancy and trophoblastic disorders

I declare that the work presented is wholly my own unless clarified as part of the submission.

I declare that no material contained in the thesis has been used in any other submission for an academic award

I confirm that the research submitted has been subject to ethical review and has not deviated from the terms of ethical approval given by the Research Ethics Committee.

Ethics ID number(s): **1766**

Signature of candidate ....Marcela B Gondek .... Date....April 2022………....

Acknowledgement

Foremost, I would like to thank my DoS Dr Frank Hills who has been a limitless source of encouragement, inspiration and intellectual drive throughout my degree. I'm certain that the high standard of academic wisdom but also humbleness and kindness set by his example will continue to guide me in my scientific career. Also his words "always under promise and over deliver" will stay with me for the years to come.

I also thank Dr Beata Burczynska, whose supervision, guidance and ability to constantly multi-task have made me believe that anything is possible. Alongside her friendliness and vivaciousness, she always provided me with continuous motivation and has been solid support along the way.

Besides; completion of this study could not have been possible without Dr Celia Bell whose kind support and belief in me, inspired me to start this bumpy journey and challenged me to grow as a scientist. I also thank my other inspiring co-workers including Dr Ajit Shah, Dr Lucy Ghali, and Dr Malcolm Ward for their insightful comments and help with all the hard questions. I wish to thank all the amazing, powerful women/mums who took part in this study and their encouragement, dedication and belief in this project that kept me determined to complete this study. There are also so many of my dearest Friends and colleagues who thought me that coffee, consistency and a good sense of humour will solve most (scientific) problems. Also, there were many others, too many to mention, who provided help and support along the way. You know who you are, thank you.

Last but not least I would like to thank my amazing Parents, my favourite Brother, the rest of my Family and my very special Dr Baker. I can't thank you enough for your consistent support throughout the years and believe in me to pursue my dream and succeed at this challenge.

List of Contents

List of Figures

List of Tables and Charts

List of abbreviations

(kilo) Dalton	(k) Da
°C	Degree Celsius
3D	Three dimensional
ACN	Acetonitrile
ActD	Actinomycin D
AMBIC	Ammonium bicarbonate
ASA	Acetylsalicylic acid
Asn	Asparagine
C	Carbamidomethyl
CBC	Complete blood count
CHM	Complete hydatidiform mole
Con A	Concanavalin A
CTP	Carboxyl-terminal peptide
CV	Coefficient variation
DHB	2,5-Dihydroxybenzoic acid
Dol-P	Dolichol head group
DTT	Dithiothreitol
EGFR	Epidermal growth factor receptor
ELISA	Enzyme-linked immunosorbent assays
ER	Endoplasmic reticulum
ETT	Epithelioid trophoblastic tumour
FCS	Fetal Calf Serum
FSH	Follicle-stimulating hormone
g/L or mg/mL	(mili) grams per (mili) litre
GC	Gestational choriocarcinoma
GlcNAc	N-acetylglucosamine
GlcNAcT-III	β1,4-N-acetylglucosaminyltransferase III
GlcNAc-Ts	N-acetylglucosaminyl transferases
GPH	Glycoprotein hormone
GTD	Gestational trophoblastic disease
GTN	Gestational trophoblastic neoplasia
h	Hour
HCCA	Alpha-Cyano-4-hydroxycinnamic acid
hCG	Human chorionic gonadotropin
hCGcf	Core fragment hCG
hCGh	Hyperglycosylated hCG
hCGn	Nicked hCG
hCGα	Free-α-subunit
hCGβ	Free- β- subunit
Hex	Hexose
HILIC	Hydrophilic Interaction Chromatography
hPL	Human placenta lactogen
HPLC	High-pressure liquid chromatography
HRP	Horseradish peroxidase
IAA	Iodoacetamide
ISOBM	International Society of Oncology and Biomarkers
LC-MS	Liquid chromatography Mass Spectrometer
LH	Luteinising hormone

LOB	Limit of blank
LOD	Limit of detection
M	Oxidation
M or mM	(mili) Mole
m/z	Mass to charge ratio
MALDI -TOF	Matrix-Assisted Laser Desorption/Ionisation Time of Flight
MS	Mass Spectrometry
Man	Mannose
MEM	Modified Eagle's Medium
min	Minutes
mIU/mL or IU/L	(mili) International units per (mili) liter
MTX	Methotrexate
n	Number
NBS	New England Bio Labs
NeuAc	Sialic acid
NGC	Non-gestational choriocarcinomas
N-linked	N-linked β-N-acetylglucosamine N-GlcNAc
O-Fuc	O-fucose
O-Gal	O-galactose
O-linked	O-linked β-N-acetylglucosamine O-GlcNAc
O-Man	O-mannose
PBS	Phosphate Buffered Saline
PGC	Porous Graphitic Carbon
PHM	Partial hydatidiform mole
PMF	Peptide mass fingerprinting
PNGase F	Peptide-N-Glycosidase F
ppGalNAc-Ts	Polypeptide N-acetylgalactosaminyl transferases
PSTT	Placental site trophoblastic tumour
PTM	Posttranslational modification
rpm	Rotation per minute
RT	Room temperature
SA	Sinapinic Acid
SD	Standard deviation
SDS-PAGE	Sodium dodecyl sulphate–polyacrylamide gel electrophoresis
SEM	Standard error of the mean
Ser	Serine
SF	Serum free
SPE	Solid-phase extraction
TFA	Trifluoroacetic acid
Thr	Threonine
TMB	3,3'5,5' Tetramethylbenzidine
TSH	Thyroid-stimulating hormone
UHQ	Nuclease-free, ultra-pure water
UV	Ultra violet

1. Introduction

In the following introductory section, the classification, molecular structure and biological functions of human chorionic gonadotropin and its isoforms are explored. Information regarding hCG glycoprotein heterogeneity and fundamental mechanisms controlling protein glycosylation and glycan composition observed in healthy pregnancy and association with malignant transformations are presented and current challenges in GTD detection and the most recent techniques with diagnostic potentials are explored.

1.1 Importance of hCG

Human chorionic gonadotropin is a heterogeneous glycoprotein hormone produced by placental trophoblasts, non-trophoblastic diseases, trophoblastic tumours and also, in relatively low concentrations, by the pituitary gland (Woldemariam and Butch, 2014). Attempts to identify its role in pregnancy considerably expanded in 1927, when the first urine test designed by Aschheim and Zondek identified the presence of a gonad-stimulating substance, which was later confirmed as hCG glycoprotein in urine samples (Lunenfeld, 2004). Since 1975, there have been many more reports on the hCG 3-D structure, its N- and O-linked glycosylation pattern, detection of its isoforms in various gestational trophoblastic diseases following many studies focused on clinical applications of hCG glycoforms reviewed by Lustbader *et al*., 1996, Cole, 2007, Kovalevskaya, 2002, Nwabuobi *et al.,* 2017. Additionally, biosynthesis, receptor activity, biological functions studies by Merz, 1994 and Koistinen *et al.,* 2015 and gene expression by Burczynska *et al*., 2013, and many more. Nevertheless, the heterogeneity and vast potential of this glycoprotein are not yet fully understood (Ibeto *et al*., 2020).

1.2 Structure of the glycoprotein hormone family

Human chorionic gonadotrophin is a member of the glycoprotein hormone (GPH) family which also comprises follicle-stimulating hormone (FSH), luteinising hormone (LH) and one non-gonadotropin member thyroid-stimulating hormone (TSH). hCG and LH are directly related in evolutionary origin. Within the same hormone family, individual forms exhibit many similar and distinctive structural and biological characteristics. For instance, the α subunit of GPHs is encoded by the same gene that is expressed in pituitary gonadotrope, thyrotrophin cells and chorionic syncytiotrophoblastic cells of all vertebrates. Thus, all GPHs consist of two non-covalently linked alpha (α) and beta (β) subunits and each subunit contains a central cysteine knot motif, which determines its 3D structure as shown in Figure 1.1 (Esteves, 2015).
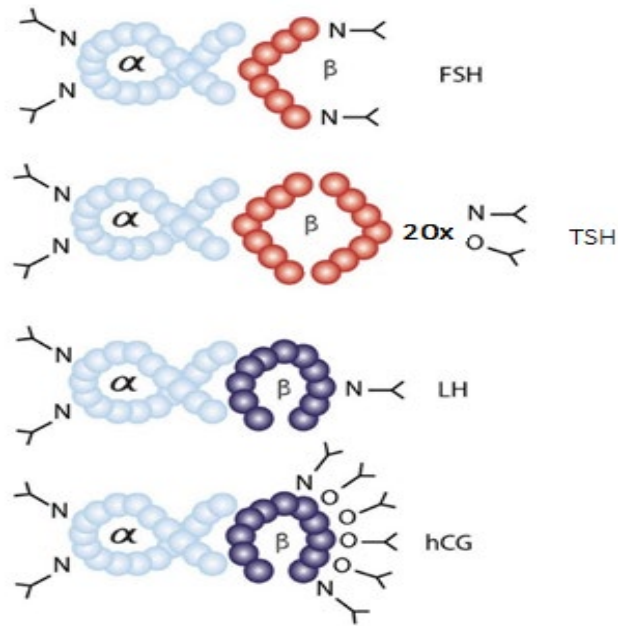
Figure 1.1 Glycoprotein hormone family structures (FSH, TSH, LH, hCG) including various N- and O-linked sugar moieties attached to alpha and beta subunits (Adapted from Esteves, 2015).

Also, the alpha subunits of FSH, TSH, LH and hCG are identical with two N-linked glycosylation sites and unique β subunits with a variable number of N-linked and O-linked glycosylation sites which confer biological specificity for each hormone (Cahoreau *et al,* 2015). Moreover, the β subunit shows ≥85% homology between hCG and LH (Figure 1.2) (Bulun, 2011, Berger and Lapthorn, 2016). Therefore, it is only due to 24 amino acids on the carboxyl-terminal end of hCGβ that the two hormones differ. hCG contains an additional 24 amino acid extension, known as carboxyl-terminal peptide (CTP) or C-terminal extension containing further potential glycosylation sites and increased structural diversity (Berger *et al*., 2013, Stenman, 2013, Esteves, 2015). It is important to mention that GPH molecular structure contributes to their maintenance in the circulation, i.e. GPHs glycoforms with CTP exhibit longer 1.5-2.5 days half-life and without a very short 5-30 min. Then, there are either eliminated from the blood by liver capture, kidney glomerular filtration or urine (Cahoreau *et al.*,2015). These findings are significantly important for the timing and accuracy of exogenous hCG measurements (pregnancy testing) (Smitz and Platteau, 2020).

Ser – Lys – Glu – Pro – Leu – Arg – Cys – Arg – Pro – Ile – Ásn – Ala – Thr – Leu – Ala – Val –
Arg   *Trp*   *His*   **13**   *Ile*

[N]

Glu – Lys –Glu – Gly – Cys – Pro – Val – Cys – Ile – Thr – Val – Ásn – Thr – Thr – Ile – Cys –
                                           30

Ala – Gly – Tyr – Cys -Pro – Thr – Met – Thr – Arg – Val – Leu – Gln – Gly – Val – Leu – Pro –
                        Met   *Ala*

Ala – Leu – Pro – Gln – Val – Val – Cys – Asn – Tyr – Arg – Asp – Val – Arg – Phe – Glu – Ser
Pro   *Thr*

– Ile – Arg – Leu – Pro – Gly – Cys – Pro – Arg – Gly – Val – Asn – Pro – Val – Val – Ser – Tyr
                                     Asp   *Phe*

– Ala – Val – Ala – Leu – Ser – Cys – Gln – Cys – Ala –Leu – Cys – Arg – Arg – Ser – Thr – Thr
-Pro   *Arg*   *Gly-Pro*
                                        Ser

– Asp – Cys – Gly – Gly – Pro – Leu – Asp – His – Pro – Leu – Thr- Cys-Asp – Asp – Pro – Arg
               Lys   *His – Pro – Gly –*

                         [O]   [O]
– Phe – Gln – Asp – Ser – Ser – Ser – Ser – Lys – Ala – Pro – Pro – Pro – Ser- Leu – Pro
- Leu - Ser– Ser– Gly – Ile – Ile – Phe – Ile (STOP)   **127**
                             121

        [O]   [O]
– Pro – Ser – Arg – Leu – Pro – Gly – Pro – Ser – Asp – Thr – Pro – Ile – Leu – Pro – Gln
    132   **138**   **145**

Figure 1.2 The amino acid sequence of LHβ compared to the hCGβ subunit. LHβ (in italics) and hCGβ (in normal text). N indicates asparagine residues with N-linked glycans, while O indicates serine residues with O-linked glycans (Adapted from Cole, 2009).

1.2.1 Structure of hCG

The hCG has a molecular weight of 36-42 kDa and contains 237 amino acids (AA) divided into two (alpha (α) and beta (β)) non-covalently bound subunits. As a protein folds, disulphide linkages are generated within the protein, which has a direct impact on its tertiary structure and role in the stabilisation of quaternary form (Valmu *et al.*, 2006, Cahoreau *et al.,* 2015). The hCG alpha subunit comprises 92 amino acids and has a molecular weight of 14.9kDa. It is folded into three loops, which are stabilized by five disulphide bonds: 7-31, 59-87, 10-60, 28-82, and 32-84 which are called the cysteine knot. Two N- linked oligosaccharides are attached to the second and third loops of asparagine residues: Asp52 and Asp78 (Figure 1.3) (Xing *et al*., 2001). The hCGβ consists of two β hairpin loops (β1 and β3) stabilized on one side of the central cysteine knot by five disulphide bonds at positions 23-72, 26-110, 34-88, 38-90, 9-57 and a long loop on the other side, forming the third- β2 loop. The segment which holds the β-subunit in place and wraps through an α subunit loop contains two β hairpin loops also known as a "seat belt" (deMedeiros and Norman, 2009). The heterodimeric subunits are orientated in an antiparallel fashion with the single loop of each subunit located nearby to the paired loops of the other subunit see Figure 1.4 below (Xing *et al.,* 2001, de Medeiros and Norman, 2009). The positioning of the two cysteine residues (93-100 and 26-110) on the outside of the hCG molecule results in them being very reactive and unstable, therefore easily undergoing reduction and losing biological activity by loss of disulphide bonds. The cysteine knot (β34-88 and β38-90) is formed and bridged by disulphide bonds through which the β9-57 residue penetrates (see Figure 1.3 – yellow lines) (Mise and Bahl, 1981, Chen *et al.,* 1992). Additionally, the hCGβ chain has a highly glycosylated carboxyl-terminal peptide (CTP) extension, with four O-linked oligosaccharides: Ser121, Ser127, Ser132 and Ser138 (see Figure 1.4) (Berger *et al.,* 2013).
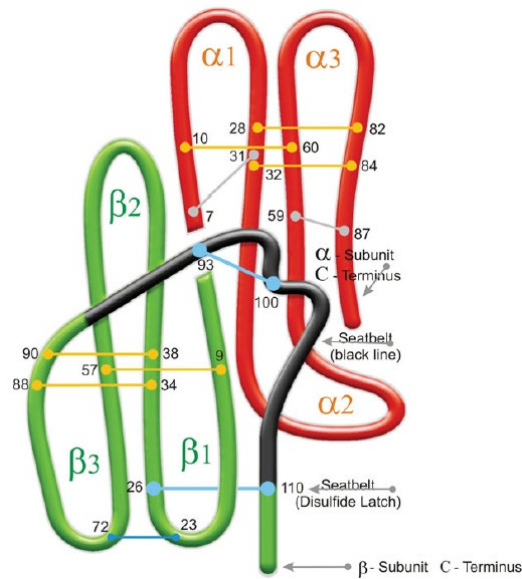
Figure 1.3 hCG structure: alpha (red) subunit, beta (green) subunit. Numbers show all 22 cysteines residues found on both subunits which are bound via 11 disulphide bridges (yellow lines- cysteine knots, grey lines– disulphide linkage on α chain and blue lines - seatbelt). Black thick lines and blue thin lines represent seatbelt (β93-110 and β25-110) and small loop disulphide bond (β 23-72), which keeps αβ heterodimer assembled (Adapted from deMedeiros and Norman, 2009).
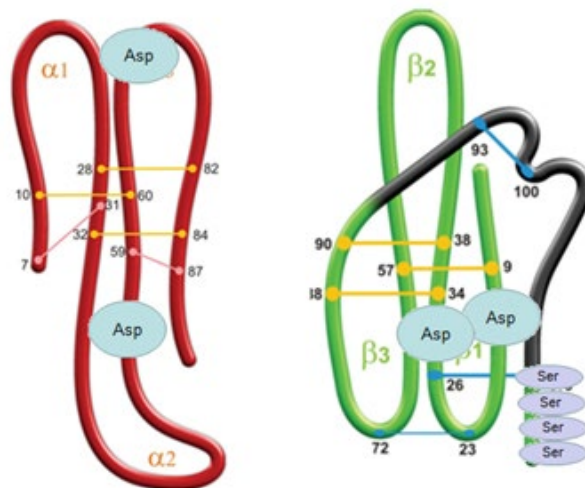
Figure 1.4 hCG subunits with marked N-linked glycosylation sites: alpha chain: Asp52 and Asp78 and beta chain: Asp13 and Asp30. Carboxyl-terminal peptide (CTP) extension (thick black line) with O-linked oligosaccharides: Ser121, Ser127, Ser132 and Ser138. Yellow and pink and blue lines represent disulphide bonds. (Adapted from deMedeiros and Norman, 2009).

1.3 Classification of hCG isoforms

Prior to secretion into circulation, newly synthesised hCG can be modified into various subunits listed below in Figure 1.5 and Table 1.1. This altered heterogeneity of the hCG molecule was primarily observed in the early 1990s by Birken and Cole's group. They detected fluctuating concentration levels of hCG isoforms, in commercially available reference standard preparation of hCG, which was purified from large pools of human pregnancy urine. That finding indicated that the hCG isoforms found in urine may differ depending on the individual and the particular week of gestation (Birken *et al.,* 1991). Supporting that discovery, in 1993 Hoermann*, et al*., showed that the serum and urinary-derived hCG, as well as tumour and pregnancy hCG, may have diverse isoform spectrum compositions. Thus, hCG molecules differ in size and charge depending on the origin (Hoermann *et al.,* 1993). Moreover, during excretion the circulating in the placenta, blood and urine intact hCG can be partially degraded into its component subunits allowing different variants of hCG to be detected in liquid samples, such as nicked hCG (Elliot *et al.,* 1997), hCG free β-subunit, hyperglycosylated hCG free β-subunit (Valmu *et al*., 2006) or core fragment hCG (hCGcf) (Stenman *et al*., 2006) (Figure 1.5). Consequently, the International Federation of Clinical Chemistry (IFCC) Working Group (WG) on Standardization of hCG Measurements has identified and defined six protein backbone variants of potential clinical interest, which are listed in Table 1.1. Due to the variable synthesis of hCG by syncytiotrophoblasts and cytotrophoblasts, levels of those isoforms fluctuate throughout pregnancy or pregnancy disorders. Moreover, alteration in hCG glycosylation will result in altered molecular stability and solubility, cellular activities such as protein folding, cell adhesion, migration and signalling, and metabolic process such as hormonal activity (Cole, 2012, Berger *et al.,* 2013). Thus, it is suggested that some forms of hyperglycosylated hCG (hCGh) and hCG free β-subunits were identified as specific biomarkers for the detection and monitoring of trophoblastic or non-trophoblastic cancers including ovarian and cervical (Cole, 1998), colon (Lundin *et al*., 2001), bladder (Mora *et al.,* 1996) and lung cancer (Yokotani *et al*., 1997).
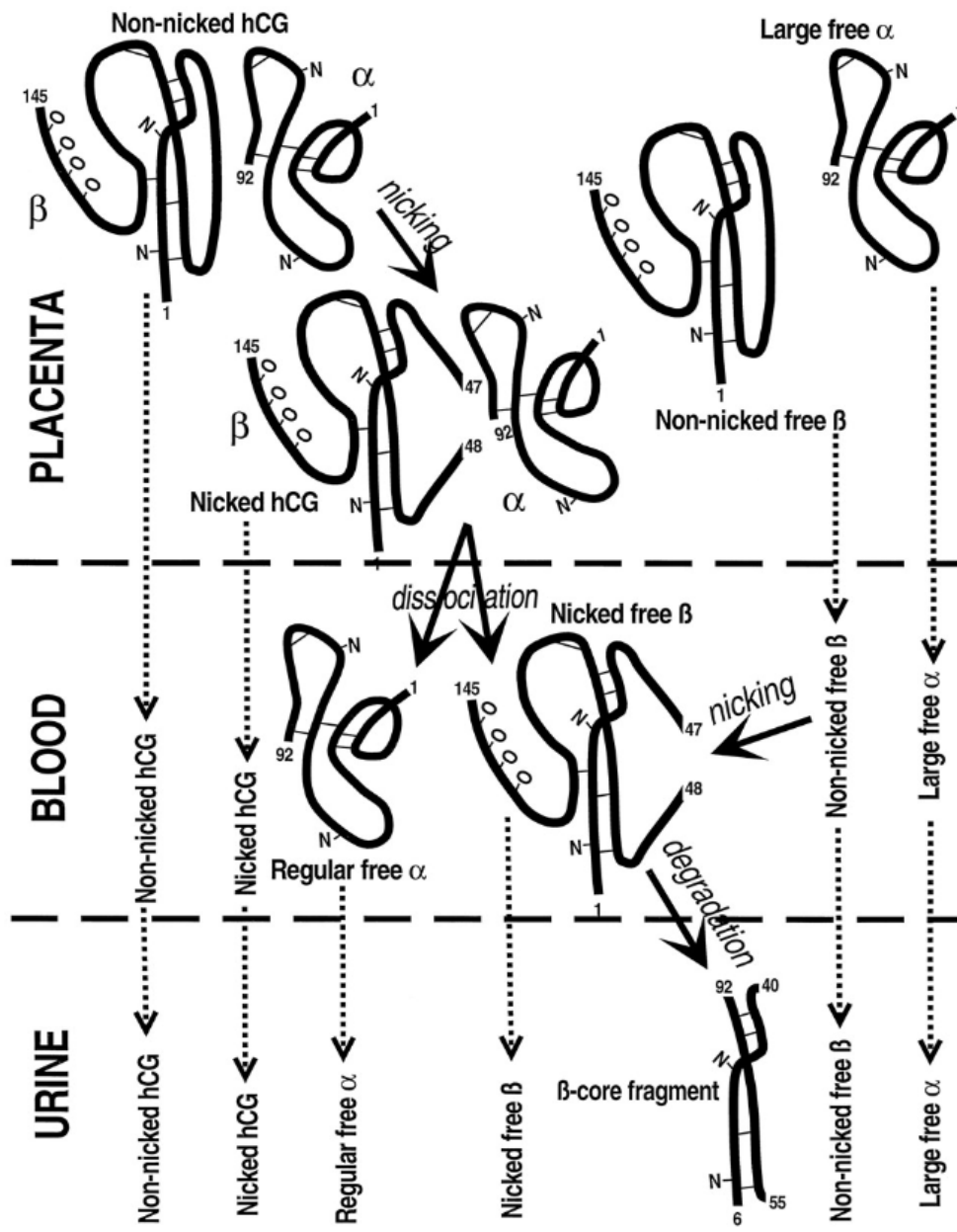
Figure 1.5 Various hCG isoforms detected in the placenta, blood, and urine (Cole, 1997).

Table 1.1 Types and descriptions of eight hCG variants. Molecular weight and sugar content need further investigation for some of the isoforms (Adapted from Cole, 2012, Berger *et al*., 2013).

Name	Symbol	Molecular weight	sugars %	Sample	Clinical use
Intact hCG	hCG	37kDa	30%	urine serum	ectopic pregnancy GTDs cancer
Nicked hCG	hCGn			plasma serum	pregnancy trophoblastic neoplasia
Free-α-subunit	hCGα	14.5kDa		plasma serum	pituitary tumours, malignant gastroenteropancreatic tumours
Free- β- subunit	hCGβ	23kDa	31%	plasma	pregnancy choriocarcinoma
Nicked free-β-subunit	hCGβn			urine plasma	Benign gestational trophoblastic disease (hydatidiform mole)
β-core fragment	hCGβcf	14kDa		urine plasma	pregnancy Down syndrome
Hyperglycosylated hCG	hCGh	43kDa	39%	serum urine	-
Hyperglycosylated hCGβ	hCGβh	28kDa	42%	serum urine	-

1.4 Genetics of hCG

As described in Section 1.2.1 hCG, like other gonadotrophic glycoproteins, is comprised of two subunits: the β subunit is expressed from a region on chromosome 19 at 19q 13.32 by a cluster of 6 hCGβ genes coding *CGB1, CGB2, CGB3, CGB5, CGB 7* and *CGB8* and a single gene located on chromosome 6q12-q21 that encodes the α subunit. In 2013, during The International Society of Oncology and Biomarkers Tissue Differentiation 7 (ISOBM TD-7) workshop, a panel of sixty-nine antibodies directed against six hCG variants: hCG, hCGn, hCGβ, hCGβn, hCGβcf and hCGα was analysed and characterised (Berger *et al.,* 2013). The assignment of diagnostically relevant antibodies to hCG epitopes has been very difficult and a comprehensive array of anti-hCG antibody:antigen complex, with accurate data on the molecular location of recognized epitopes, is still not accessible. It is caused by the reduced accuracy of comprehensive arrays affected by hCG crystallization and protein degradation. Additionally, the majority of epitopes on hCG, continuously overlap each other, and it is very challenging to draw definite borders between them and arrange structurally characteristic antigenic domains (Berger, 2013, Berger and Lapthorn, 2016). Thus a fixed number of antigenic domains on hCG and its variants has been determined and the number of epitopes within these antigenic domains ultimately corresponds to the theoretical antibodies repertoire that is generated (Lustbader *et al.,* 1989, Berger *et al*., 2002). Therefore, the majority of epitopes are located in two major antigenic domains: the tips of β sheet loops 1 and 3 (epitopes β2-6) and around the cystine knot (epitopes β1, β7 and β10), also on hCGβcf (epitopes β11-13) and on the hCGβ- CTP (β8 and β9) (Figure 1.6). Five epitopes have been defined on hCGα and some of these are exposed only on the free β subunit while others are exposed both on the free subunit and the intact heterodimeric hormone (Berger, 2013, Stenman *et al.,* 2013). Subsequently, harmonization of antigenic domains and epitopes is crucial for reliable and reproducible hCG measurements. For the purpose of pregnancy tests, non-selective detection of hCG variants could be enough, however further separation of hCG and hCGβ measurements can help to distinguish between normal pregnancy and GTDs, such as choriocarcinoma. Furthermore for other diagnoses, such as testicular cancer or non-trophoblastic diseases, the selectivity of the tests has to increase and provide distinctive separation and ratio of various hCG isoforms. Therefore, tests that will focus on detecting, purifying and characterizing hCG glycoforms will improve current diagnostic methods and therefore patient management (Ibeto *et al*., 2021).
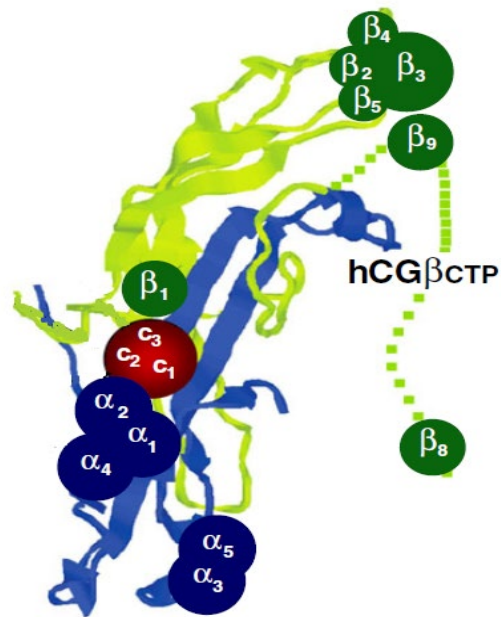
Figure 1.6 Map of hCG epitopes. α epitopes (presented in blue), β epitopes (shown in green), and epitopes determined by the central cystine knot-based epitope cluster (shown in red). (Adapted from Berger *et al*., 2013).

1.5 Glycobiology of hCG glycoforms

Glycans are not encoded directly from DNA sequences but assembled through a series of individually catalysed reactions. A few per cent of known genes produce the enzymes and transporters, responsible for the biosynthesis and assembly of glycans typically as posttranslational modifications (PTMs) of proteins (Varki and Kornfeld, 2017). One of the most complex forms of PTMs is called glycosylation, with numerous possibilities of glycan assembly and composition. The attachment of sugar moieties to proteins by various glycosidic linkages, including N- and O-linked glycosylation, occurs in the endoplasmic reticulum (ER) and Golgi apparatus. It is determined by the translation and transcription of genes coding for a variety of competing and sequentially acting glycosidases and glycosyltransferases (GlcNAc-Ts) which expand the functional diversity and increase the structural heterogeneity of proteins (Taylor and Drickmer, 2011).

Figure 1.7 shows a brief overview of glycoproteins synthesis from protein translation to circulation. In eukaryotes synthesis of polypeptides and protein folding takes place on the surface of ER. Then the growing peptide chain is shuttled into the lumen of ER, where the core glycan processing is performed by various glycosyltransferases. Next in the Golgi, further N- and O- glycosylation takes place following, glycoproteins packing into secretory vesicles bound for fusion with the plasma membrane. Finally, secreted proteins are released into the extracellular space from where, they can enter circulation (Kuzmanov *et al*., 2013).
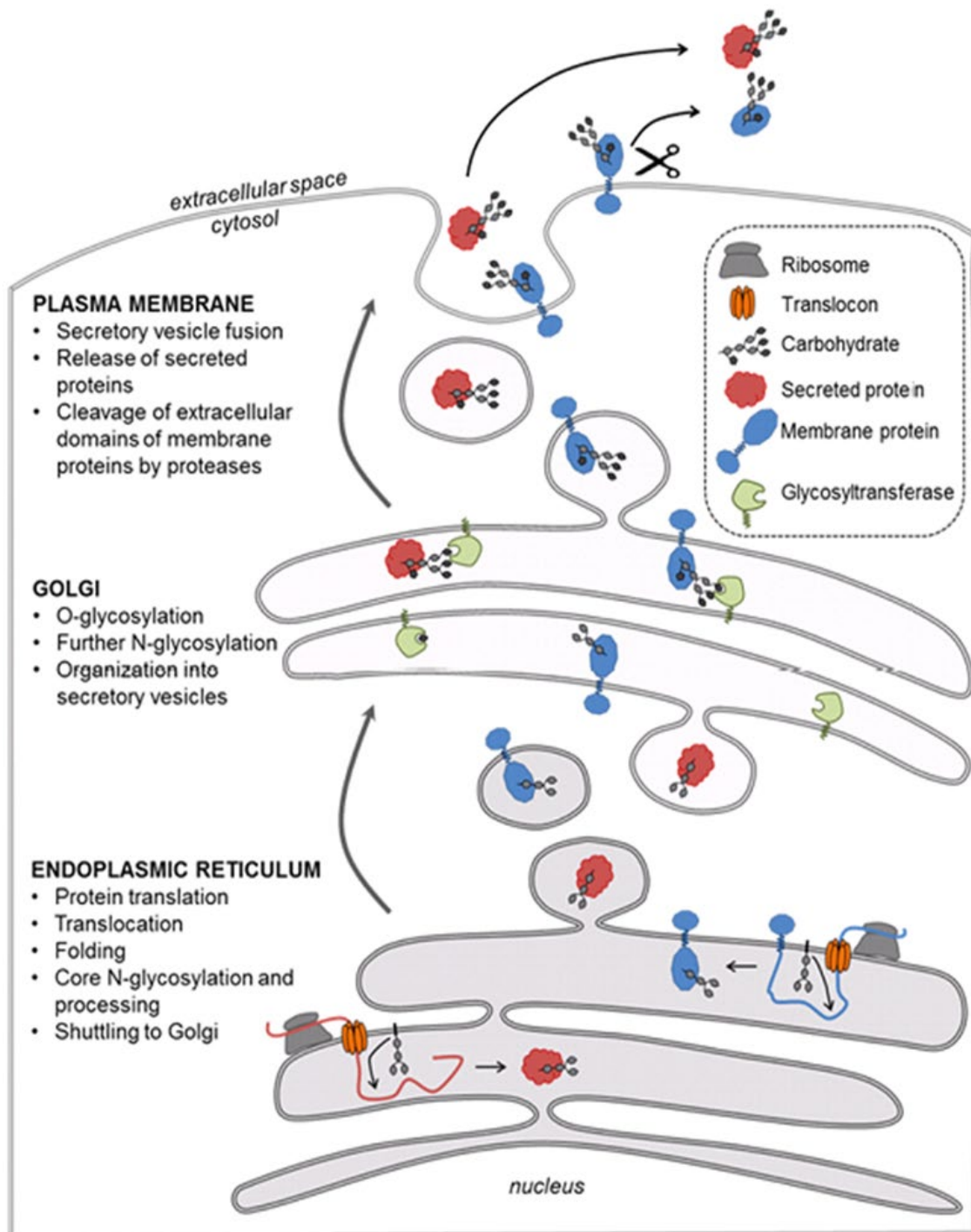
Figure 1.7 Overview of the glycoprotein biosynthesis pathway. Glycan assembled on the ER membrane and transferred to the protein. Proteins are exported to Golgi where they are processed by glycosidases and glycosyltransferases to remove or add residues. Secreted glycoproteins are released into extracellular space to enter circulation (Kuzmanov *et al.,* 2013).

The glycome is characteristic of a particular cell type at a specific state of differentiation (Nwabuobi *et al.,* 2017) therefore, even with full knowledge of the expression levels of all relevant gene products, small changes in environmental cues can cause accurate prediction of glycans compositions very challenging (Stanley *et al*., 2009). For instance, the glycosylation pattern and profile of hCG isoforms change throughout the gestation period in normal pregnancy and in pathological diseases, which will be discussed further in this project. Thus such variation in glycan structures indicates that some glycans sequences do not have universal roles in all tissues and cell types in which they are expressed (Stanley *et al*., 2009).

1.5.1 N-linked glycan assembly

As more structures are being discovered the ever-increasing structural diversity of glycosylation arises from remodelling in the ER and Golgi via variation in trimming, extension and branching by various building blocks (Gagneux *et al.,* 2017). Figure 1.8 shows an assembly of N-glycan which starts immediately after the transfer of the initial seven N-glycan residues onto the lipid dolichol head group (Dol-P) in the cytoplasmic side of the ER. Next, the entire structure is flipped to face the luminal compartment of ER and further sugar residues are added via glycosyltransferases (GlcNAc-Ts). This formation of lipid-linked precursors comprising fourteen oligosaccharides ($Glc_3Man_9GlcNAc_2$) linked via N-acetylglucosamine (GlcNAc) i.e. amide nitrogen (N-glycosidic bond) at the side chain of asparagine (Taylor and Drickamer, 2011).
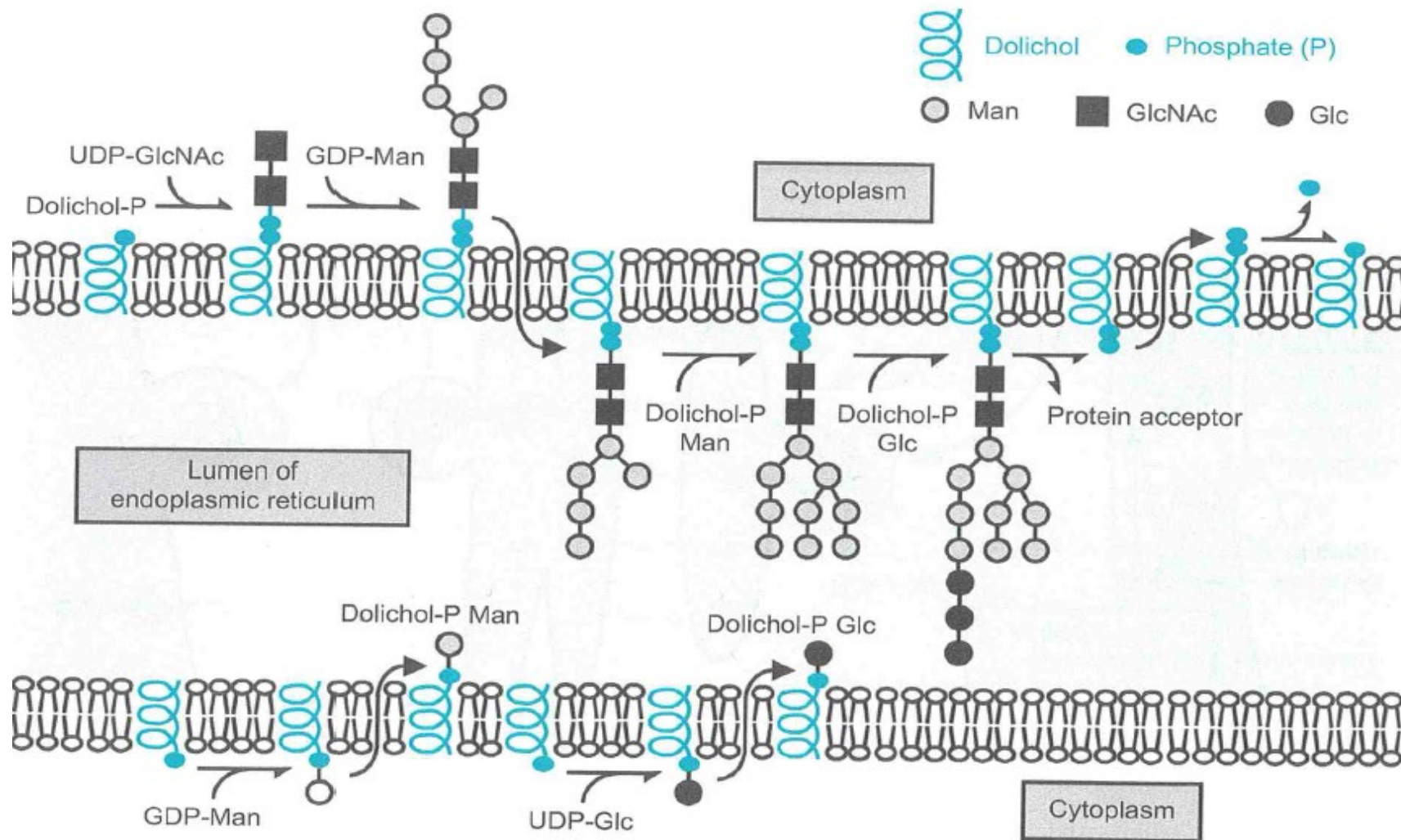
Figure 1.8 Assembly of N-linked glycans. Generation of oligosaccharide structure linked onto the lipid dolichol head group (Dol-P) (Adapted from Taylor and Drickamer, 2011).

In eukaryotes, this N-linkage commonly involves N-acetylglucosamine (GlcNAc) residue and the minimal amino acid consensus sequence of asparagine followed by any amino acid except proline (X) and serine (Ser) or threonine (Thr) i.e. Asn-X-Ser/Thr (see Figure 1.9).
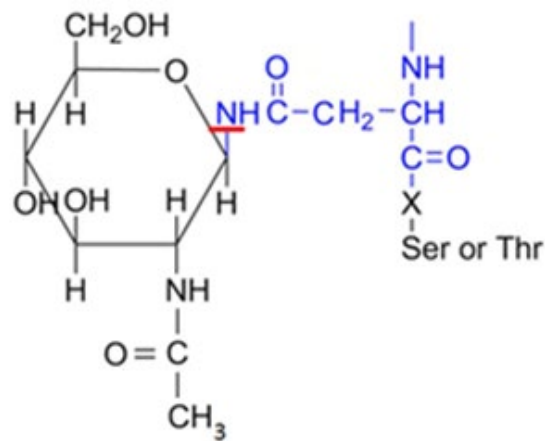
GlcNAc + Asparagine

Figure 1.9 The N-linked carbohydrates (GlcNAc) (black structure) are covalently linked (underlined red) to the amino acid asparagine (Asn) (blue structure). The N-linked sequence is Asn -X- Ser/Thr (Adapted from Varki and Lowe, 2009).

Subsequently, in the lumen ER and then in the Golgi apparatus, the complex N-glycoproteins such as hCG are modified by hydrolytic removal or addition of sugars such as galactose, fucose or sialic acid (Pinho and Reis, 2015, Jansen *et al*., 2016, Azevedo *et al.,* 2018). Identified on hCG glycosylation and high sialic acid content influence glycoprotein folding, subunit assembly and receptor binding activity (Lis and Sharon, 1993, Stowell *et al*., 2015).

1.5.2 Glycan synthesis and composition

As part of the normal human developmental programme, many terminal GlcNAc residues in N-glycosylation reactions are regulated by GlcNAc-Ts. This leads to the synthesis of complex N-glycan branching during embryogenesis and the postnatal period. Numerous acceptors and/or precursors such as galactose, GlcNAc, fucose, and sialic acid are linked to the mono-antennary pentasaccharide core. The presence of numerous GlcNAc antennas (branches) in N-linked glycans generated progressively by GlcNAc-Ts found in the Golgi apparatus can be divided into three types presented in Figure 1.10 (Kizuka and Taniguchi, 2018). Firstly high-mannose (A), which only mannose (Man) residues extend the common pentasaccharide core region containing Man3GlcNAc2Asn. Secondly, hybrid (B), in which Man extends the Manα1-6 arm of the core and 1or 2 GlcNAcs extend the Manα1-3. Finally, the complex (C), in which antennas extend the core can be classified as biantennary, triantennary and tetrantennary structures (Stanley *et al*., 2017).
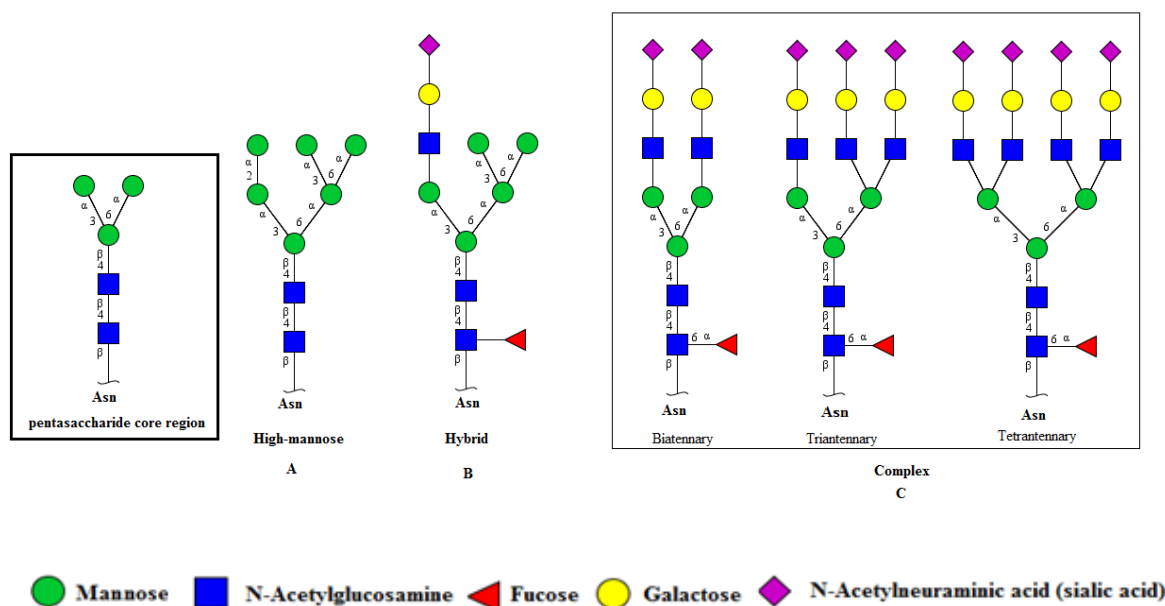
Figure 1.10 N-linked glycans composition. Pentasaccharide core common is all structures. A- High-mannose, B- Hybrid and C- Complex, including biantennary, triantennary and tetrantennary structures (Adapted from Schmidt, 2017).

In addition, a unique modification of the GlcNAc linked to the core mannose in hybrid or complex N-glycans catalyzed by β1,4-N-acetylglucosaminyltransferase III (GlcNAcT-III), termed the bisected GlcNAc is commonly found in humans (Figure 1.11). Also, glycans with sulphated terminal β-linked GalNAc are found on the pituitary glycoprotein hormones i.e. LH and TSH, but not on FSH. The synthesis of sulphated GalNAc attached to GlcNAc residues is controlled by sulphotransferases (Miwa *et al*., 2012).
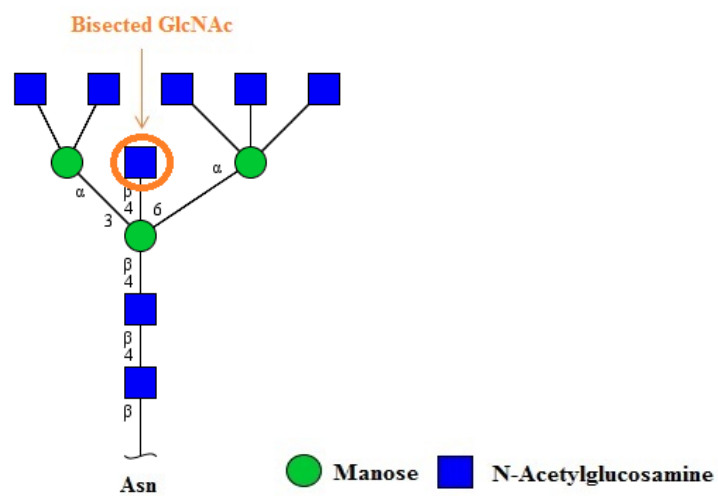
Figure 1.11 Structure of bisecting GlcNAc (Adapted from Stanley and Cummings, 2017).

In contrast, O-linked glycans are linked to serine or threonine (Ser/Thr) residues (Figure 1.12). O-linked glycosylation is initiated via attachment of N-acetyl-galactosamine (GalNAc), forming O-GlaNAc. Subsequently, O-mannose (O-Man), O-fucose (O-Fuc), O-galactose (O-Gal), or the nucleocytoplasmic glycan O-linked β-N-acetylglucosamine (O-GlcNAc) may be added (Figure 1.13). The transfer of GalNAc from the sugar donor UDP-GalNAc to Ser or Thr is controlled by UDPGalNAc polypeptide N-acetylgalactosaminyl transferases (ppGalNAc-Ts) (Reis *et al*, 2010, Pinho and Reis, 2015). In this thesis, O-linked glycosylation is not further discussed.
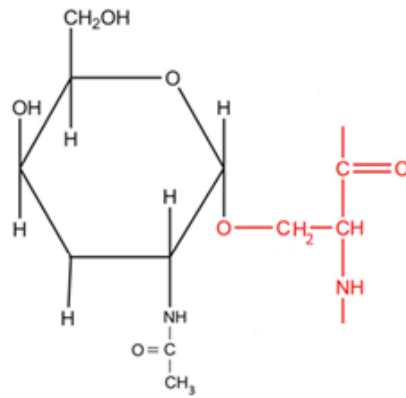
GlcNAc + Serine

Figure 1.12 O-linked carbohydrates are linked covalently through N- Acetyl-galactosamine and the amino acids serine or threonine (Adapted from Varki and Lowe, 2009).
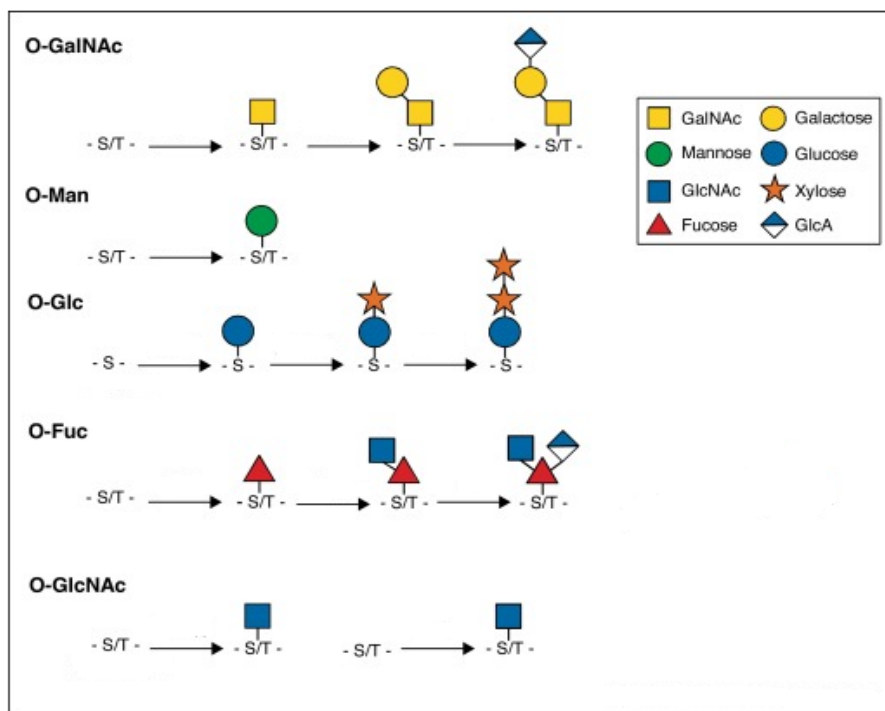
Figure 1.13 Major types of various O-linked attachments. Symbols denoting each saccharide are shown in the legend (Adapted from Zhang and Hagen, 2019).

1.6 Hyperglycosylated hCGh

The search to identify the structure of hyperglycosylated hCGh began in 1983, when Mizuochi *et al*., (1983) used hydrazinolysis, gel filtration with glycosidases and methylation analysis to show distinctive weight differences in hCG derived from choriocarcinoma (40 kDa) when compared to hCG present in pregnancy (37 kDa). That variation was studied by Pierce and Parsons (1981) who claimed that about one-third of the hCG molecular weight found in pregnancy was due to glycosylation consisting of 8 oligosaccharide chains covalently bound to the alpha and beta subunits. Later, it was reported that hCGh produced in choriocarcinoma and testicular germ cells malignancies, contains a larger amount of O-linked biantennary oligosaccharides than hCGh found in the first trimester of normal pregnancy urine, which explained some of the molecular weight variations (Cole, 1987). Further studies confirmed that the amount of and/or location of oligosaccharide chains present on the beta subunit of hCG isoforms significantly increases the molecular weight of this glycoprotein and its variation depends on the isoforms origin (Elliott *et al.,* 1997). Consequently, the term hyperglycosylated hCG (hCGh) was used to describe hCG with higher molecular weight derived from choriocarcinoma and regular hCG reserved for the hCG with normal oligosaccharide chains found during pregnancy (Cole *et al*., 1998).

All four N-linked oligosaccharides found on hyperglycosylated hCG contain an N-acetylglucosamine (GlcNAc) residue linked to asparagine Asn52 and Asn78 on the alpha chain, and Asn13 and Asn30 on the beta chain, followed by another GlcNAc, mannose and two more branches of mannose. As previously mentioned this structure has been described as a mono-antennary pentasaccharide core with the remaining components being variable (Iles, *et al.,* 2014). Moving forward, knowledge about hCG glycosylation sites carries on expanding and in 2012, an analysis of hyperglycosylated hCGh provided new details about attached glycans. It was revealed that hCG variants are the most glycosylated of glycoproteins, with hCG carrying 30% sugar by molecular weight, hyperglycosylated hCG containing 39% sugar, and hyperglycosylated hCG free ß subunit-containing 42% sugar (Cole, 2012). More recently, bisected glycans (GlcNAc) were detected on hCG molecules. These N-linked type residues constituted tri- and tetra- antennary structures and were detected at much higher levels in early and late pregnancy than in gestational trophoblastic disease (GTD) (Ibeto *et al.,* 2020). Additionally bisecting GlcNAc of N-glycans were characterised as regulators of cellular signalling and tumour progression (Miwa *et al.,* 2012) therefore, it confirms earlier suggestions that hyperglycosylated hCG (hCGh) promotes the growth of cytotrophoblasts in pregnancy and

implantation, and initiates malignant progression in choriocarcinoma (Cole *et al*., 2006). It was also suggested that it works through the inhibition of apoptosis in cytotrophoblasts and promotes cell growth (Sasaki *et al.,* 2008). Consequently, the recognition and characterisation of hCG's isoforms glycome in trophoblastic diseases such as choriocarcinoma are discussed with reference to its value as potential biomarkers.

1.7 Glycosylation in cancer cells – biomarkers potential

Glycosylation of proteins is a well-known hallmark of tumour progressions and represents a valuable source of information. Malignant transformations are associated with changes in the expression of glycan moieties and glycosyltransferase activity, which expand the degree of pre-existing microheterogeniety of individual proteins (Dennis *et al.,* 1987, Brockhausen, 2006). Changes in terminal glycan structures observed in association with malignant transformations include accumulation of high-mannose glycans due to premature termination of glycan processing, loss or excessive expression of certain glycans such as increased fucosylation (addition of fucose) and sialylation (addition of sialic acid), increased expression of incomplete or truncated (trimmed) glycans, reduction in bisecting glycans and increased branching (bi-, tri-, tetra- antennary) (Varki *et al.,* 2017, Oliveria-Ferrer *et al.,* 2017). Thus, studies for instance on epithelial ovarian cancer cell lines (OvCa) showed up-regulation of high-mannose and complex glycans, whereas downregulation of bisecting glycans in metastatic cells (Zhang *et al.,* 2019); increased levels of tri- and tetra-antennary structures and sialylation was detected in highly metastatic prostate cancer cell lines (Lange *et al.,* 2012) and breast cancer cells (i.e. MCF7) were characterized by significant increases in high-mannose and complex tri-antennary N-glycans compared to normal breast tissue (Liu *et al.,* 2013). As a result, the close relationship between hyperglycosylated protein and cancer can be employed as a biomarker to accurately detect pathogenic alterations (Varki and Lowe, 2009).

Glycoprotein cancer biomarkers already widely used and approved by the Food and Drug Administration (FDA) include α-Fetoprotein, Prostate-specific antigen, Cancer antigen 125, Human epididymis protein 4, β-2 Microglobulin + CA 125II, apolipoprotein A1+prealbumin+ transferrin, HE4 + CA125, Cancer antigen 15-3, Cancer antigen 27-29, Carbohydrate antigen 19-9 or cancer antigen 19-9, Carcinoembryonic antigen, Human epidermal growth factor receptor 2 and Thyroglobulin. (Kirwan *et al.,* 2015). These glycoprotein-biomarkers (Table 1.2) have the capability to provide minimally invasive, and accurate information about the

condition, which can be prognostic and predictive. However, only total protein levels, measured using immunoassays, are clinically monitored and used for diagnostic purposes which may indicate some limitations (Kirwan *et al*., 2015).

Table 1.2: Glycoprotein cancer biomarkers list, used in clinical practice, with The Food and Drug Administration (FDA) year of approval (Adapted from Kirwan *et al.,* 2015).

Marker	Full name	Cancer type	Year of FDA approval
AFP	α-Fetoprotein	Liver	1992/2008
PSA, Pro2PSA	Prostate-specific antigen	Prostate	1986/1994/2012
CA125 (MUC16)	Cancer antigen 125	Ovarian	1997/2011
HE4 (WFDC2)	Human epididymis protein 4	Ovarian	2008
OVA1 test (multiple proteins)	β-2 Microglobulin + CA 125II (up), apolipoprotein A1 + prealbumin + transferrin (down)	Ovarian	2009
ROMA test	HE4 + CA125	Ovarian	2011
CA15-3 (MUC1)	Cancer antigen 15-3	Breast	1997
CA27-29	Cancer antigen 27-29	Breast	2002
CA19-9	Carbohydrate antigen 19-9 or cancer antigen 19-9	Pancreatic, Ovarian	2002
CEA	Carcinoembryonic antigen	Colon, Gastric, Pancreatic, Lung, and Breast	1985
HER2/neu	Human epidermal growth factor receptor 2	Breast	1998
Tg	Thyroglobulin	Thyroid	1997
hCG	Human chorionic gonadotropin	Testicular, Ovarian	NOT APPROVED

Therefore, the development and refinement of mass spectrometric techniques, used in conjunction with the traditional methods may provide a precise and accurate understanding of the heterogeneity and complexity of protein glycosylation. This leads to the conclusion that mass spectrometric methods developed to characterise and measure complex carbohydrate structures in the response to cancer-associated stimuli may be valuable for high throughput clinical analysis.

hCG is already used as a marker for invasive trophoblast cells function and it plays a key role in the detection of ovarian and cervical malignancies (Cole, 1998, Handschuh *et al.,* 2007), colon (Lundin *et al*., 2001), bladder (Mora *et al.,* 1996), lung cancer (Yokotani *et al*., 1997), Down syndrome pregnancy (Alldred *et al*., 2015) and other pregnancy complications or pregnancy loss (Kovalevskaya *et al*., 2002). Considering that glycosylation patterns of the same protein can differ between tissues and between normal and transformed cells, the ability to detect and characterise these variations could enable a large number of glycoproteins to have tissue/tumour-specific profiles. Therefore, regardless of glycan heterogeneity (structures do not follow a predefined template), the use of sensitive analytical mass spectrometry techniques such as Matrix-Assisted Laser Desorption/Ionisation (MALDI) Time of Flight (TOF) Mass Spectrometry (MS) and Q Exactive™ Plus Hybrid Quadrupole-Orbitrap™ Liquid Chromatography-Mass Spectrometer (LC-MS) for precise glycome mapping will be discussed (Chapter 2). Consequently the ability to perform such tasks consistently and routinely will significantly widen a spectrum of potential biomarkers and therapeutic applications.

1.8 Classification of gestational trophoblastic diseases

Gestational trophoblast diseases comprise a spectrum of pregnancy-associated growths derived from the placenta; from the pre-malignant conditions of complete and partial hydatidiform mole (CHM and PHM), which might develop into gestational trophoblastic neoplasia (GTN), through to the malignant invasive mole, choriocarcinoma, placental site trophoblastic tumour (PSTT) or epithelioid trophoblastic tumour (ETT) (Seckl *et al.,* 2010).

1.8.1 Genetics of GTDs

The hydatidiform moles are associated with abnormal gametogenesis and fertilization, often following molar pregnancy; they are locally invasive and lack the tendency to develop widespread metastases. On ultrasonographic scans complete hydatidiform moles appear as a

heterogeneous mass "bunch of grapes" or "snowstorm", without associated fetal development (Seckl *et al.*, 2010). Nevertheless, choriocarcinoma is more aggressive and can follow molar pregnancy, spontaneous abortion, ectopic or normal pregnancy. Complete hydatidiform moles are usually diploids, androgenetic in origin, where empty ovum (no maternal DNA) is fertilised by a sperm, karyotype 46XX or 46XY originating from duplication of haploid sperm (Figure 1.14). A maternal autosomal-recessive missense gene mutation, most commonly *NLRP7* or *KHDC3L* on chromosome 19q 13.3-13.4, is linked to complete moles, leading to chronic molar pregnancies (Lurain, 2010, Ning *et al.,* 2019).

Partial hydatidiform moles are almost always triploid and result from the fertilisation of an ovum by two sperm cells resulting in karyotype 69XXX, 69XXY or 69XYY. PHMs have focal oedema, denatured areas of varying size and form, and pathological trophoblast cell proliferation in placental villi. There may be fetal tissue or a visible embryo present, however living infant is extremely rare (Ning *et al.,* 2019).
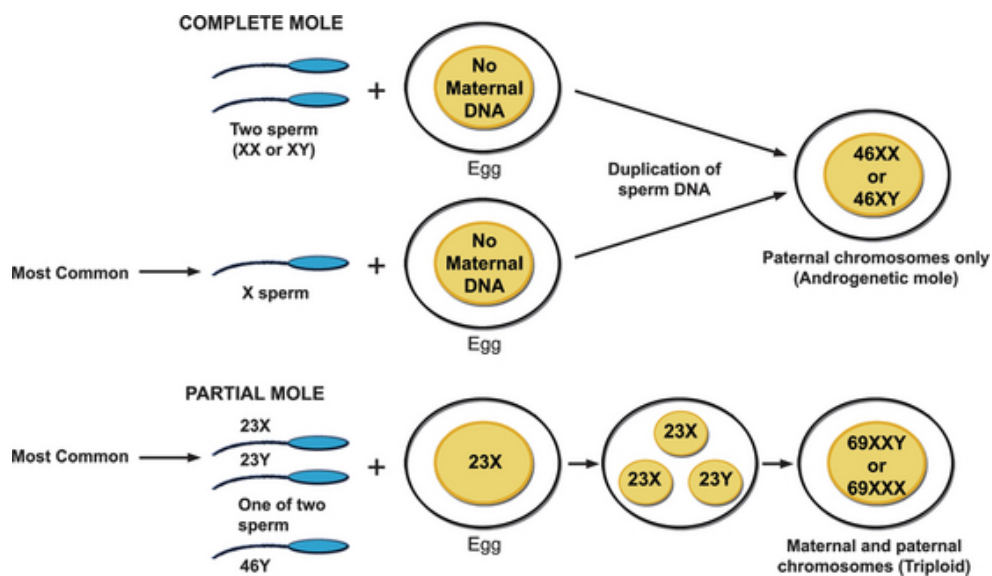
Figure 1.14 Genetic origins of molar pregnancy. Dispermic or monospermic CHM arise from fertilisation of an empty ovum with two sperms or duplication of paternal chromosomes (46XX or 46XY). Dispermic or monospermic PHM arise as a result of fertilisation of the single, viable ovum by two sperms. These triploid conceptions may be 69XXY or 69XXX (Adapted from Crum, 2005).

1.8.2 Physiology of trophoblast

In a healthy pregnancy, the monolayer of fetal trophoblast cells starts to invade the decidualized endometrial epithelium, essential to facilitate embryo implantation into the uterus. Placental development begins when the blastocysts, containing an outer layer of trophoblast cells, attach to and starts to penetrate, the endometrial epithelium. Activated trophoblast cells start differentiation and ensure normal growth and further development of the fetus. During implantation, cytotrophoblast cells fuse to form an outer layer of multinucleated syncytiotrophoblasts, which secrete endocrine and autocrine hormones such as hCG (Figure 1.15). The primary role of hCG is to maintain corpus luteum progesterone synthesis and secretion until 10-12 weeks of gestation when the placenta itself can synthesise significant amounts of progesterone. It is also an important autocrine and paracrine regulator of epidermal growth factor, transforming growth factor, and leukaemic inhibitory factor for increasing placental syncytium formation and in blastocyst implantation i.e. it suppresses the maternal immune system toward fetal and placental tissues (Licht *et al.,* 2007, Cole, 2009).
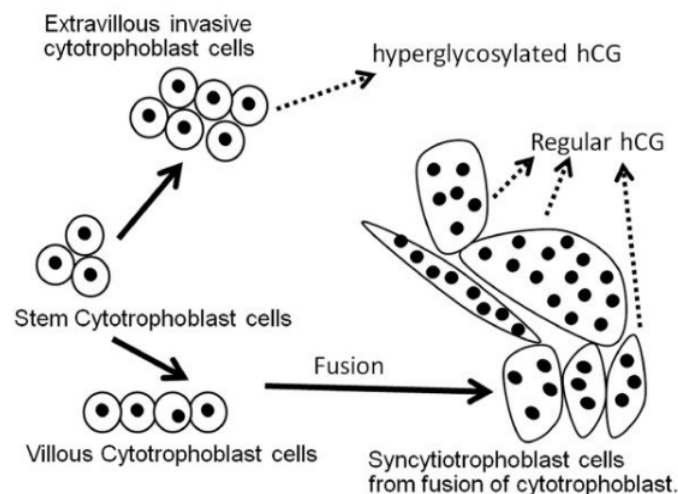
Figure 1.15 Representation of trophoblast cells differentiation in placental villi and hCG secretion (Adapted from Cole, 2009).

hCG mediates glycogenolysis in human placental villi, stimulates prostaglandin synthesis by placental tissue and inhibits myometrium contractility during pregnancy which promotes fetal organ growth (Demers *et al*., 1973). In addition, hCG regulates the vascular tone in cord blood and amniotic fluid and attenuates the vascular response to angiotensin II (Demers *et al*., 1973,

Ticconi *et al*., 2007). The hCGβ is mainly connected to inhibitory or stimulatory activity in cellular growth and hCGα is linked with prolactin secretion and control of endometrial cell differentiation (Blithe *et al.,* 1991, Gillot *et al.,* 1996, Hermsteiner *et al.,* 2002, Rao and Lei, 2007, Cole, 2009). In maternal serum, hCG can be detected eight days post ovulation (Braunstein *et al*, 1976) and in the blastocysts seven days after fertilization (Hay and Lopata, *et al*., 1988). Gradual increase of hCG can be detected until peak levels are reached at 8 to 12 weeks. During the first semester, levels increase two-fold every 48 hours; during the second trimester, the levels gradually decrease to plateau and are maintained during the third trimester until decrease around week 20 (Figure 1.16). Additionally, hCG concentration can vary daily in the maternal serum (Lapthorn *et al*., 1994, Wide *et al*., 1994, Kovalevskaya *et al.,* 2002, Guibourdenche *et al.,* 2010). About 2 weeks after implantation extravillous cytotrophoblasts invade further into the decidualized endometrium, which is essential for successful pregnancy (Kliman, 2009, Cole, 2009).
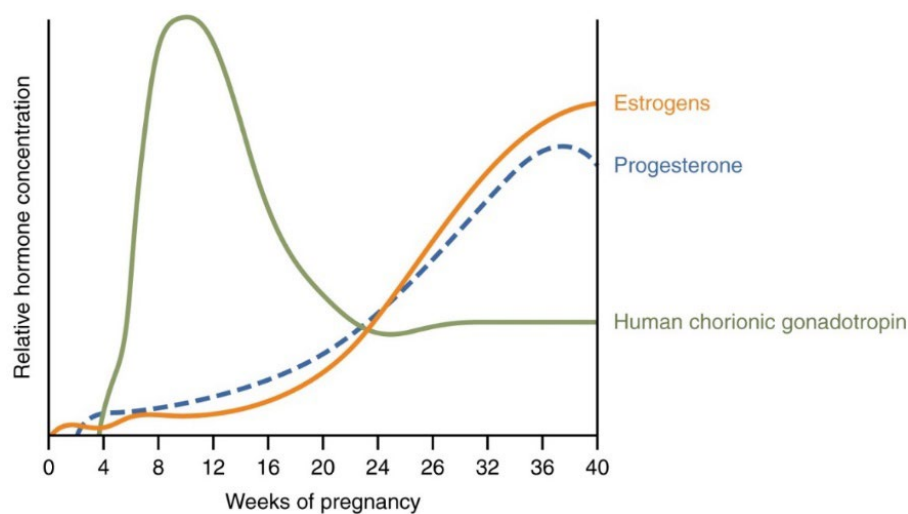
Figure 1.16 Rates of secretion of progesterone, estrogen and hCG at different stages during pregnancy (OpenStax, 2013).

Any disturbance in these processes is associated with pregnancy complications such as pre-eclampsia, ectopic pregnancy, and gestational trophoblast diseases. Therefore, pre-eclampsia is detected in 2-5% of pregnant women and it is the most common cause of maternal and fetal morbidity and mortality. Patients experience hypertension, proteinuria and oedema in the

second and third trimesters of pregnancy. An increased mitotic activity stimulates cellular proliferation in the cytotrophoblast and elevation of hCG isoforms in serum and urine samples. A low concentration of hyperglycosylated hCGh in the first trimester is caused by syncytiotrophoblasts undergoing cellular necrosis. Treatment with acetylsalicylic acid (ASA) is recommended to reduce the risk of preeclampsia (Roberge *et al*., 2012, de Medeiros and Norman, 2009, Stenman *et al*., 2013).

Ectopic pregnancy is the consequence of a fertilised ovum being implanted outside the endometrial lining of the uterus and is associated with low levels of hCG during early pregnancy. Symptoms include low abdominal pain and vaginal bleeding. Early diagnosis has been enhanced by improving serum hCG assays, in conjunction with high-resolution ultrasound, and increased awareness and understanding of the associated risk factors when using assisted reproductive technology (ART). Methotrexate may be used as an alternative to surgery to resorb ectopic pregnancy. It is suggested that low hCG levels (<5000 IU/L) are a good indication of the success of this treatment (Sivalingam *et al*., 2012).

GTDs including hydatidiform moles and choriocarcinoma arise from villous trophoblast tumours with abnormal villous structure, hyperplasia, stromal karyorrhectic debris, stromal hypercellularity and collapsed villous blood vessels. Meanwhile, early PHM shows abnormally-shaped irregular villi, trophoblastic pseudo inclusions and patchy trophoblast hyperplasia (Seckl, 2009). An invasive mole has a similar histological presentation as a complete mole, however, it is characterized by myometrial invasion driven by hyperglycosylated hCGh, without intervening in the endometrial stroma. Invasive moles consist of a significant amount of hCGh (30±35%) of total hCG, which distinguishes them from non-invasive moles (Cole, 2009). After the evacuation of a molar pregnancy, the patient has usually presented with heavy vaginal bleeding and it can metastases to the lung, vagina, cervix, vulva and gastrointestinal tract (Murdoch *et al.,* 2006, Drife and Magowan, 2005).

Lastly, PSTT is very rare (≤2% of GTD), mostly occurs after pregnancy, however, it can also arise after abortion, miscarriage or hydatidiform mole. Complete removal of PSTT is only achieved by surgery as the American Cancer Society warns that these forms of GTDs are very resistant to chemotherapy drugs. Additionally, 15-20% of PSTT might become malignant with metastases to the lungs, liver, abdominal cavity and brain (Figure 12). Comparatively low hCG levels and high human placental lactogen levels are found in the serum of patients (ACS, 2017).

1.8.3 Gestational trophoblastic neoplasia (GTN)

Choriocarcinoma is a highly aggressive, malignant neoplasm which originates from abnormal proliferation of atypical villous trophoblasts cells without the development of villi. A common subtype is thought to derive from cytotrophoblastic cells (Figure 1.17) that have undergone a neoplastic transformation during a molar pregnancy (androgenetic subtype) (a) or in a non-molar placenta (bi-parental subtype) (b) is gestational choriocarcinoma (GC). For the purpose of this research, we will be only focusing on GC, however, alternative subtypes which are non-gestational choriocarcinomas (NGC) arise from germ cells (c) or somatic tumours (d) (Xing *et al.*, 2019). Accurate discrimination between these subtypes is vital for clinical management as non-gestational choriocarcinoma can be less responsive to chemotherapy. Approximately 50% of choriocarcinomas arise from complete hydatidiform moles (CHM) and a complete molar gestation, around 30% occur after a miscarriage or ectopic pregnancy and 20% can be found after normal pregnancy (even after 10-15 years) (Zhang *et al.*, 2019).
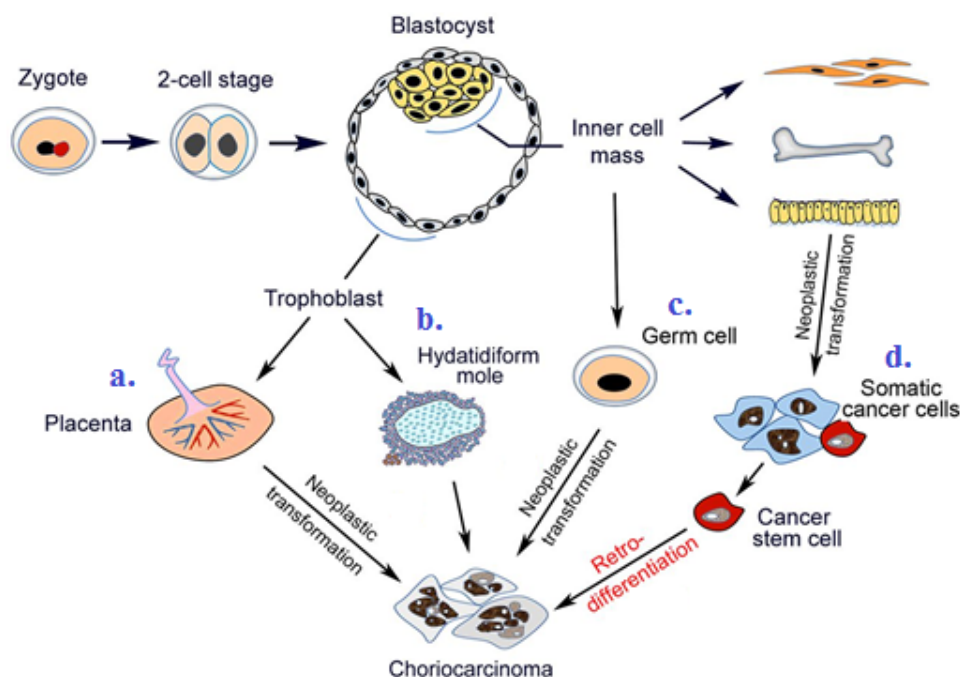
Figure 1.17 Development of gestational and non-gestational choriocarcinoma. Gestational choriocarcinoma originated from cytotrophoblastic stem cells in (a) placenta or (b) hydatidiform mole. Non-gestational choriocarcinoma of (c) germ cell origin and retro-differentiation of (d) somatic cancer stem cells (Adapted from Xing *et al.,* 2019).

The GC and NGC exhibit similar morphological patterns, histopathological classification and biochemical markers. Many women with choriocarcinoma will present with shortness of breath (dyspnoea), anaemia, pre-eclampsia, hyperemesis, hyperthyroidism, neurological symptoms and abdominal pain, additionally due to elevation of hCG levels abnormal uterine bleeding can be observed. Furthermore, as mentioned above it might develop in areas other than the uterus and not be related to pregnancy. It can occur in both men and women and spread to ovaries, testicles, chest or abdomen. Such type of choriocarcinoma is usually linked with other types of cancer, which form mixed germ cell tumours (Zhang *et al*., 2019).

Therefore the importance of the distinction between therapeutic response and prognosis of the patients is clinically significant. Survival for patients with NGC, following surgical tumour resection and multidrug treatment with several chemotherapy agents, is very low, around 16% (Mello *et al.,* 2017). In contrast, patients with GC such as choriocarcinoma treated with chemotherapy and without hysterectomy have a much higher survival rate of nearly 90% (Li *et al*., 2015). On the other hand, in male malignancies, testicular germ cell tumours account for only 1% of solid tumours in reproductive men, however, survival decreases from 90% in non-seminomatous germ cell tumours to 48% in aggressive, poor-prognosis non-seminomatous choriocarcinoma (Rejlekova *et al*., 2019).

1.8.3.1 Gestational Choriocarcinoma

Molar pregnancy, induced and spontaneous abortion, ectopic pregnancy, and term or pre-term deliveries are all common causes of gestational choriocarcinoma in women of reproductive age. The incidence is estimated at 1 in 40 000 - 50 000 pregnancies and 1 in 40 hydatidiform mole cases (Mello *et al.*, 2017).

Figure 1.18 below, illustrates malformed placentas, with the abnormal proliferation of villous trophoblasts caused by genetic aberrations. The numerous vesicles, representing complete hydatidiform mole (A) display acute swelling (hydropic degeneration) of all the chorionic villi filling the uterus. Conversely in a partial hydatidiform mole, a live fetus associated with healthy chorionic villi (C) or abnormally shaped irregular, patchy trophoblast hyperplasia and no healthy fetus (B) with interspersed molar vesicles can be observed. Additionally in invasive hydatidiform mole, molar vesicles penetrate the myometrium giving rise to a mass distorting the uterine wall (D). Further invasion of trophoblastic cells, in the absence of chorionic villus structures, may extend into and beyond the myometrium – choriocarcinoma (E). Prognosis of

choriocarcinoma is strongly impacted by cell necrosis, haemorrhage, vascular invasion and remote metastases especially to vagina, spleen and kidney, gastrointestinal truck, lung (E) or brain and liver. Lastly, PSTT is composed of intermediate trophoblastic cells from the extravillous trophoblast and invades the myometrium by separating muscle bundles and fibres (F) (Cooper, 2016, Vuong *et al.,* 2000).

A. Complete hydatidiform mole   B. Partial hydatidiform mole   C. Coexistent mole & live fetus

D. Invasive mole   E. Choriocarcinoma   F. Placental site trophoblastic tumour
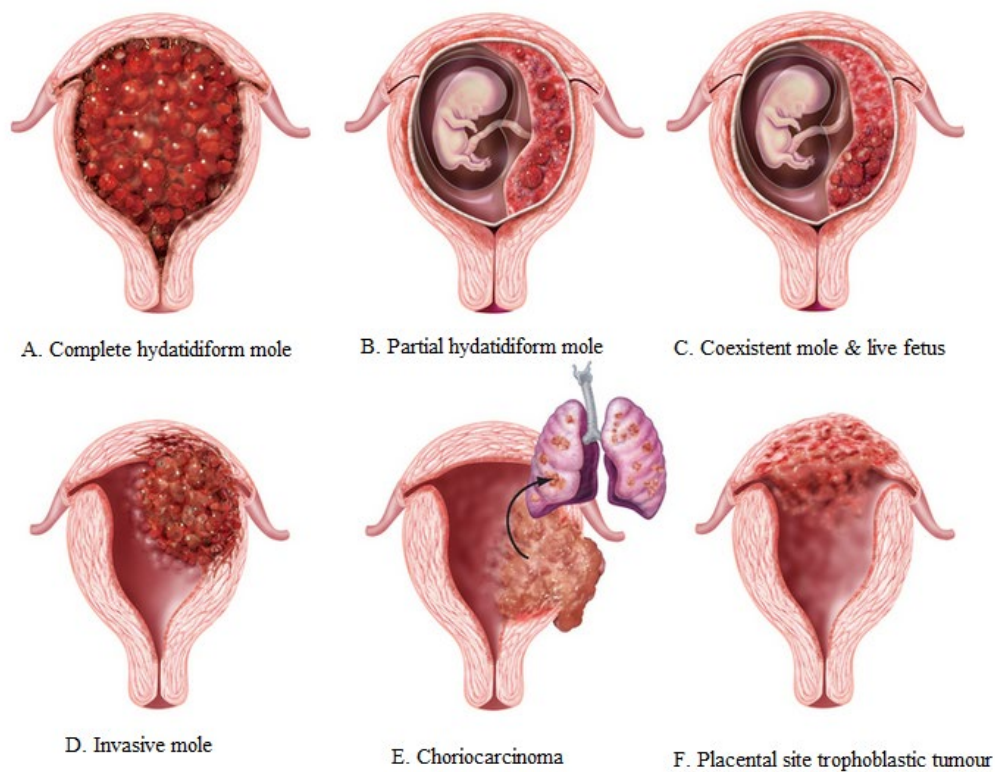
Figure 1.18 Model of various gestational trophoblastic neoplasia (A, B, C) and persistent gestational trophoblastic disease (D, E, F) (Adapted from Cooper J.A., 2016)

1.8.3.2 Detection and treatment of choriocarcinoma

Investigations for clinicopathologic features of choriocarcinoma include complete blood count (CBC) with platelets, levels of AE1/AE3 (cytokeratin used to identify epithelial cells), hPL (human placenta lactogen), Ki-67 (cell proliferation marker), PLAP (placental alkaline), phosphatase coagulation studies, renal function panel and liver function panel. Moreover, the age of the patient (reproductive-postmenopausal) is analysed. Also, tumour classification (GC or NGC), tumour location, metastasis (vagina, uterine cervix, liver, lungs, brain) are investigated and the number of previous pregnancies, number of live births, history of pregnancy (hydatidiform mole, endometrial curettage, term pregnancy). Lastly, the period between last pregnancy and beginning of treatment, prognostic score based on the International Federation of Gynaecologists and Obstetricians (FIGO), prognostic score risk ranked into low risk scoring 0-6 and high risk with score >6 and quantitative levels of total hCG (mIU/mL). Some of those tests are invasive, time-consuming and might contain false positive or false negative results (Lurain, 2011, Mello *et al.,* 2017). However elevated levels of human placental lactogen hormone (hPL) and chorionic gonadotropic hormone (hCGβ), especially detected after pregnancy, are used for diagnostic purposes and they account for 75% of invasive mole and 25% of choriocarcinoma cases (Vuong *et al*., 2000, Arabi *et al*., 2021). Additionally, ultrasound and histopathological investigation of haematoxylin and eosin-stained tumour tissues or immunostaining for p57 are commonly used (Lurain, 2010, Xing *et al.,* 2019).

To date, the most validated method for early choriocarcinoma diagnosis is hCG monitoring i.e. rising or plateauing levels of hCG isoforms defined as; a rise in days 1,7, 14, and 21 over three weeks or two consecutive rises in days 1,7 and 14 over two weeks determine patient therapy (Seckl *et al*., 2010, Bishop and Edemekong, 2019). During chemotherapy, hCG levels are measured at least once per week and conclusive follow-up after molar evacuation requires measurements being taken between 1-2 weeks until 3 consecutive tests show normal levels. After that levels should be measured two weekly (during the first year) four weekly (after year 2), eight weekly (year 3), 3-monthly (year 4), 4-monthly (year 5) or 6-monthly (after year 5) for life, due to the risk of relapse after chemotherapy (Seckl *et al.,* 2010). In the United Kingdom patients start chemotherapy when serum total hCG levels are greater than 20.000 IU/L, one month after HM evacuation or it is vaginal or lung metastasis greater than 2 cm (Chan *et al*., 2021). Interestingly, structural variations in hCG isoforms led to a lack of consensus regarding, optimal, sensitive and qualitative analytical methods for its detection. Commercially available assays/pregnancy tests will detect and measure intact hCG and hCGh

isoform simultaneously and give the positive result, using the combination of hCG proteins as a pregnancy biomarker (Cole, 1998). Unfortunately, these results lack precision and accuracy and can cause false-positive results. Moreover, immunoassays do not measure the bioactivity of hCG but the concentration of epitopes recognised by the antibodies used in the immunoassays, hence even degraded and not bioactive forms are detected, giving misleading results. The ratio between immunoreactivity and bioactivity of hCG changes with the type of new standards and immunoassays. Thus, there is a clear need to use faster more effective, accurate and precise separation and detection techniques for hCG isoforms, that can be used in diagnostic and therapeutic procedures (Berger *et al*., 2013). Furthermore, available treatment options for choriocarcinoma include chemotherapeutic agents such as methotrexate (MTX), calcium folinate (folinic acid), actinomycin D (ActD), etoposide, cyclophosphamide, doxorubicin, melphalan, hydroxyurea, vincristine or cisplatin, dilatation and curettage, hysterectomy or combination of all (Ngan *et al.,* 2021). In summary, choriocarcinoma is a very rare malignant disease that is usually associated with a gestational trophoblastic disease, such as hydatidiform mole, spontaneous abortion, or ectopic pregnancy. Currently, the diagnosis includes a physical examination followed by an ultrasound scan and a thorough check of the person's medical history, including current and previous pregnancies, vaginal bleeding, headache, cough, vomiting, purple/blue nodules in the vagina with enlarged uterus size. Also, blood or urine samples are collected and tested for hCG levels (βhCG, hCGβ-cf, C-terminal hCG, and preferably hyperglycosylated hCG), as if a person has raised hCG in their blood and urine but is not pregnant, this may be a sign of GTDs such as choriocarcinoma. That will trigger an advanced and prolonged (2-3 weeks) medical investigation, including measuring of hCG levels and histological examinations of the patient's biopsy (Ngan, et al., 2021). Subsequently, the tendency of choriocarcinoma for early vascular invasiveness accounts for the high risk of metastasis, which mostly affects the lung (80–85%) but also affects the vagina, brain, kidney, stomach and small bowel (Gasparri, et al., 2019). Therefore, precise and expedited differential diagnoses (both with single or multiple metastases) are essential in choriocarcinoma care. As surgery (dilatation and curettage or hysterectomy) and differential diagnosis obtained after tumour removal may cause patient distress and unnecessary treatment including chemotherapy (hCG levels >20 000 mIU/mL) i.e. EMA-CO Treatment (Etoposide-Methotrexate-Actinomycin (Dactinomycin)-Cyclophosphamide-ONCOVIN). As a result, our initiative is to provide a glyco-biomarker that may specifically predict choriocarcinoma at the earlier stage. This will potentially help provide these patients with earlier and more accurate diagnoses while also improving their quality of life.

1.9 Purpose of work described in this thesis

As discussed previously to reliably diagnose and treat gestational trophoblastic diseases such as choriocarcinoma, it is crucially important to precisely and accurately detect and measure hCG glycoforms levels in biological samples. Early diagnosis of cancer allows a good prognosis for the patient and at present, the availability of screening and diagnostics tests for choriocarcinoma, which include invasive and time-consuming examinations and measurement of human placental lactogen hormone (hPL) and chorionic gonadotropic hormone (hCGβ) is not at a satisfactory level and can be improved. Also, the paucity of measurement tools for precise identification and quantitation of various hCG glycoforms is still a challenge and therefore the discovery of new biomarkers, which alone or in combination with existing techniques can be applied to improve earlier diagnosis and treatment of choriocarcinoma is advisable. Therefore, observations of established glycoprotein pathways in cancer, the characterization of hCG glycoforms and the development of robust techniques may lead to the development of pioneering, reliable biomarkers. The current explosion of proteomic and glycomic data in the recent disease-specific research provides new recommendations and equally new challenges in clinical diagnostics. However, there is a clear indication based on previous and ongoing studies at Middlesex University and not only, that by optimising and enhancing current proteomics and glycomics techniques; the development of new powerful screening tests for detection of altered glycosylation in neat urine samples is achievable and can be used as a diagnostic tool.

1.9.1 Hypothesis

We hypothesise that hCG glycoforms characterised from healthy pregnancy urine and clinical samples of choriocarcinoma urine will differ in glycan amount, composition and structure and that these patterns will be easily distinguished using mass spectrometric tools. Therefore, all these changes could form the basis for a novel glyco-biomarker discovery marker.

1.9.2 Research aims and objectives

The overall aim of this project is to compare the heterogeneity of hCG glycoforms present in normal pregnancy and choriocarcinoma

The specific aims of the project are:

I.   **Purify hCG from biological samples**

In order to achieve this we will:

➢ Optimise culture conditions for maximising hCG production from trophoblastic cell lines
➢ Use a panel of hCG antibodies to purify hCG isoforms from biological samples including choriocarcinoma cell lines, normal placental cell lines, choriocarcinoma urine and normal pregnancy urine and assess the efficiency of developed techniques

II.   **Develop and optimise a mass spectrometry-based method for the identification and detection of carbohydrate moieties on human chorionic gonadotropin glycoforms**

In order to do this we will:

➢ Optimise methods for deglycosylation, extraction of glycans, detection and characterisation using MALDI-TOF MS and GlyoQuest protein scape database
➢ Compare heterogeneity of hCG glycoforms throughout normal pregnancy and choriocarcinoma

III.   **Characterization of hCG glycoforms from the biological samples listed in objectives from aim 2.**

In order to achieve this, we will:

➢ Use the methods developed from aims 1 and 2 to determine the composition of isolated N-linked glycans using GlycoQuest software
➢ Use information from our studies and from existing glycan databases to develop a database for hCG N-linked glycans

2. Development of methodology for detection and enrichment of hCG in complex biological fluids

This chapter describes the development of a method for detecting and isolating human chorionic gonadotropin (hCG) from a variety of biological materials. Firstly, for the determination of typical hCG levels in human pregnancy urine, commercially available hCG standards i.e. AC048 (Fitzgerald, USA) and C0434 (Sigma, UK) were tested. They were derived from urine collected during all pregnancy trimesters and pooled together from a healthy pregnant individual/s and are commonly used for routine diagnostic measurements of hCG (Nwabuobi, *et al.,* 2017). Certificates of the analysis of used standards confirmed high purity ≥ 98% SDS-PAGE. A spectrum of hCG variants i.e. intact hCG, hCG β subunit and hCG α subunit were identified in these preparations, and both standards served as a source to study hCG glycosylation. Additionally, as a promising model for the study of hCG derived from primary human syncytiotrophoblasts, conditioned media of human trophoblast cells isolated from human choriocarcinoma cells were used: JEG-3, (BeWo, JAR - data not shown). Furthermore, as a guide for hCG levels and glycosylation diversity in urine during normal pregnancy and gestational trophoblastic disease, human pregnancy urine from seven healthy females, collected between 2016 to 2019 in England, UK and five clinical urine samples from choriocarcinoma patients, collected between 2014-2016 at Medical University Hospital in Poznan, Poland were tested. Pregnancy urine samples from typically developing pregnancies (week 4 to 33 of gestation) were collected, with informed consent and institutional ethical approval from Natural Sciences Sub-Committee (NSESC) at Middlesex University - Approval for project no 1766. For the collection of choriocarcinoma samples, the urine collection protocol of the Human Tissue Act and Material Transfer Agreement was followed.

Initially, the procedure used for hCG isoforms purification from biological samples included sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) however, this method was not sufficiently separating proteins from urine samples and could not accurately detect low concentrations of hCG forms suitable for further glycan analysis. Hence, in order to purify hCG from biological samples such as urine and conditioned media, three enrichment methodologies were investigated. These comprised of binding hCG according to their carbohydrate-binding properties (lectin) or by using solid phase-conjugated antibodies (protein G sepharose and paramagnetic beads). Additionally, peptide mass fingerprinting (PMF) identification technique using Matrix-Assisted Laser Desorption /Ionisation (MALDI) Time of Flight (TOF) Mass Spectrometry (MS) was acquired to accurately and rapidly validate all

detected hCG isoforms. Moreover, hCG isoforms concentration in all samples was measured using enzyme-linked immunosorbent assays (ELISA). Consequently, a reproducible analytical method for hCG purification and detection, utilising solid-phase extraction combined with MALDI–TOF MS analysis was developed and applied to human pregnancy and choriocarcinoma urine.

2.1 Materials and Reagents

2.1.1 General major and minor equipment

Centrifuge	Sigma Aldrich, UK
pH Meter	ThermoFisher, UK
Heating block	ThermoFisher, UK
Balance	ThermoFisher, UK
Vacuum concentrator	Eppendorf, UK
Vortex mixer	ThermoFisher, UK
Pipettes	Brand, Germany
Magnetic stirrer	ThermoFisher, UK
NanoDrop spectrophotometer	ThermoFisher, UK
Magnetic rack	ThermoFisher, UK
Thermomixer	Eppendorf, UK
MALDI TOF MS/MS (Autoflex)	Bruker, UK
CO_2 incubator	Binder, Germany
SDS-PAGE tank and gels	Bio-Rad Laboratories, UK
MTP 384 ground steel ID210 1000476	Bruker, UK

2.1.2 General chemicals

- Phosphate Buffered Saline (PBS) tablets, BR0014, Oxoid, UK; formula: 8g sodium chloride, 0.2g potassium chloride, 1.15g disodium hydrogen phosphate, 0.2g potassium dihydrogen phosphate were dissolved in nuclease-free water, pH7.3

- Trifluoroacetic acid (TFA) 99+%, for HPLC (10294110), Fisher Chemical, ThermoFisher, UK

- Nuclease-free water (10429224), UHQ, Invitrogen™, ThermoFisher, UK

- 50 mM Ammonium bicarbonate (AMBIC) was prepared by adding 79mg of AMBIC (10207183) which is 99% pure powder used for analytical analysis purchased from ACROS Organics™, (ThermoFisher), UK to 20ml of Nuclease-Free Water (10429224), UHQ, (Invitrogen™ThermoFisher, UK)

- 0.1% trifluoroacetic acid (TFA- T62200), Sigma Aldrich, UK, was prepared by adding 200μL of TFA in 200mL of distilled water

- Acetonitrile (ACN) for HPLC-MS (10616653), ThermoFisher, UK

2.1.3 hCG preparations

- hCG collection from culture media: human placental, choriocarcinoma cells lines JEG-3 (HTB-36, ATTC, UK), BeWo (CCL-98, ATTC, UK), JAR (HTB-144, ATTC, UK) and extravillous trophoblast cell line HTR-8 (CRL-3271, ATCC, UK), cultured in Dulbecco's Modified Eagle's Medium (MEM) with 0% or 10% Fetal Calf Serum (SF or 10% FCS) (Sigma, UK) and Insulin, Transferrin and Selenium (ITS) supplement (ThermoFisher, UK), all with 1% Streptomycin/Penicillin (Sigma, UK)

- hCG collection from human pregnancy urine of seven healthy participants (4-33 weeks) collected between November 2016 to March 2019

- Urine samples from choriocarcinoma patients, collected between January 2010 to April 2013 (Wielkopolskie Oncology Center, Poznan, Poland)

- Commercially available hCG standards purified from human pregnancy urine: intact hCG (AC048), Lot A14072402, Fitzgerald Laboratories, USA and intact hCG (C0434, Batch 1 and 2), Lot SLCB3890 and Lot SLCB3785 Sigma, UK

2.1.4 Reagents for hCG detection

2.1.4.1 Trypsination and Sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE)

- Dithiothreitol (DTT - D0632) ≥98% (HPLC), ≥99.0% (titration), Sigma Aldrich, UK

- Iodoacetamide (IAA - A3221) ≥99% (HPLC), Sigma Aldrich, UK

- Trypsin from porcine pancreas, Sigma Aldrich, USA, Cat No T6567, Lot 031M6160V, for MALDI analysis was prepared by reconstituting the lyophilized product in 50 mM Ammonium bicarbonate and 10% acetonitrile (stock solution 100ng/µl)

- Precision Plus Protein Standards Dual Colour Ladder, 161-0394, Bio-Rad, UK

- 4–15% Mini-PROTEAN® TGX (4561083SEDU), BioRad, UK, precast polyacrylamide gels with a 10-well comb

- Sample buffer (2x Laemmli, dilute 1:1 with sample) 62.5 mM Tris-HCl, pH 6.8, 2% SDS, 25% glycerol, 0.01% bromophenol blue

- Running buffer (Tris/glycine/SDS) 25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3

2.1.4.2 Reagents for Enzyme-linked immunosorbent assay (ELISA)

- Human hCG beta DuoSet ELISA, DY9034-05, R&D Systems, UK

  Kit contains ELISA plate coating buffer (896036), reagent diluent concentrate (841380), colour reagent A and B (895000, 895001), stop solution (895926), wash buffer (895003)

- hCG ELISA kit, SKU: HC251F, Calbiotech, USA

  Kit contains microwell coated with Streptavidin plate, hCG standards, hCG conjugate, TMB substrate, stop solution and 20x wash concentrate solution

2.1.4.3 Reagents for Matrix-Assisted Laser Desorption/Ionisation (MALDI) TOF MS/MS

- 2,5-Dihydroxybenzoic acid (2,5 DHB- 85707) matrix substance for MALDI-MS, >99.0% (HPLC), Sigma Aldrich, UK

- Alpha-Cyano-4-hydroxycinnamic acid (4-HCCA-70990) matrix substance for MALDI-MS, 99.0% (HPLC), Sigma Aldrich, UK

- Sinapinic Acid (SA- D7927) matrix substance for MALDI-MS, ≥99.0% (T)

2.1.5 Reagents for hCG enrichment

- Anti-hCG beta 5008 SP-5 antibody (1mg/ml), Product code 100006, Medix Biochemica (MedixMAB™), Finland. Specificity for intact hCG molecule and its free β subunit

- Anti-hCG beta 5004 SP-1 antibody (1mg/ml), Product code , Medix Biochemica (MedixMAB™), Finland. Specificity for intact hCG molecule and its free β subunit

- Concanavalin A (Con A) (05-0106), Medicago AB, Uppsala, Sweden

- Protein G Sepharose (CL-4B), GE Healthcare, Sigma Aldrich, UK

- Dynabeads® M-280 Tosylactivated Cat No 14203, Invitrogen by Thermo Fisher Scientific, Lot 00753595

- Dynabeads Buffer B: 0.1M sodium phosphate buffer, pH 7.4 was prepared by adding 52.4mg of sodium phosphate monobasic (NaH_2PO_4 - S0751, Sigma Aldrich) and 288.4mg of sodium phosphate dibasic (Na_2HPO_4 – S9763, Sigma Aldrich) in 20mL nuclease-free water.

- Dynabeads Buffer C: 3M ammonium sulphate buffer was prepared by adding 396.4mg of ammonium sulphate ($(NH_4)_2SO_4$ - A4915, Sigma Aldrich) in 1 mL of 0.1M sodium phosphate buffer

- Dynabeads Buffer D: PBS with 0.005% Tween 20 (15825398, Fisher Scientific) was prepared by dissolving 1 PBS tablet, Oxoid, UK in 100mL nuclease-free water and adding 5µL Tween 20

- Dynabeads Buffer E: PBS with 0.001% Tween 20 (15825398, Fisher Scientific) was prepared by dissolving 1 PBS tablet, Oxoid, UK in 100mL nuclease-free water and adding 1µL Tween 20

- 13mM ammonia solution, pH 10.3 was prepared by adding 144µL of 32% ammonia (1054261000, Supelco, Sigma Aldrich, UK) into 200mL of distilled water

2.2 Methods

2.2.1 Preparation of hCG

Commercially available hCG standards, hCG isolated from cultured human trophoblast cells, from urine collected during normal pregnancy and from choriocarcinoma patients were tested.

2.2.1.1 Commercially available hCG standard

The commercially available standards/controls of hCG; intact hCG (AC048, 1 mg/ml), Fitzgerald Laboratories, USA and intact hCG (C0434, 1mg/mL), Sigma, UK were used. Both standards were purified from healthy pregnant urine pooled from the 1st, 2nd and 3rd trimesters of pregnancy. They were reconstituted in nuclease-free water (ThermoFisher, UK), and aliquoted (1-100μL) at concentrations (0.3 -20μg/mL), then kept at -80° C to minimise degradation.

2.2.1.2 hCG secretion from trophoblast cells

To detect and measure hCG secreted into media from human placental, trophoblast cells; extravillous trophoblast HTR-8 and choriocarcinoma JEG-3 cell lines ($12x10^6$ in 175 cm^2 flasks) were cultured in serum-free (SF) Dulbecco's Modified Eagle's Medium (MEM) for 42 days with supplementation (around 20%) of fresh media. In order to maximise hCG yield, human placental cells lines JEG-3, BeWo and JAR were grown at different cells densities (8 $x10^6$ and 12 $x10^6$ cells in 75 cm^2 and 175 cm^2 tissue flasks) for up to 42 days. In addition, to increase hCG secretion, cells were cultured in a variety of media including Dulbecco's Modified Eagle's Medium (MEM) with 0% or 10% Fetal Calf Serum (SF or 10% FCS) and Insulin, Transferrin and Selenium (ITS) supplement, all with 1% Streptomycin/Penicillin, following incubation at 37° C, 5% CO2 for 21 days (not all data presented). In each case, a sample of cell-free conditioned media was collected from fully confluent flasks and stored at -80° C for further analysis. hCG was detected using SDS-PAGE and concentrations were determined using an hCG enzyme-linked immunosorbent assay ELISA kit, Calbiotech, USA.

2.2.1.3 Collection of urine samples with hCG

Urine collection, processing, storage and experimental protocols were approved by Ethics Committee at Middlesex University - Natural Sciences Ethics Sub-Committee (NSESC) - Application for Ethical Approval (Rev 01 Oct 13, 1766) and performed in accordance with relevant guidelines and regulations. Informed consent was obtained from all participants in this study and urine samples were fully anonymized. Morning urine samples from seven healthy

pregnant women were collected between 2016 and 2019, subjects at various gestational ages were collected in sterile containers (100-500 mL) and centrifuged within 1 hour of collection at 5000 rpm for 10 min. The gestational age was confirmed by subjects with reference to medical history. Sediment included cells and cell debris were disposed and clarified supernatant was immediately stored at -80° C for further analysis. Urine from cancer patients (10-20mL) was collected between 2014 and 2016, after diagnosis of choriocarcinoma (performed at the Medical Science University, Poznan) and before surgical intervention and chemotherapy. Urine was collected in sterile containers and centrifuged within 1 hour of collection at 5000 rpm for 10 min. Sediment included cells and cell debris were disposed and clarified supernatant was immediately stored at -80° C for further analysis. 1mL of urine from five patients was transferred to Middlesex University for further analysis. hCG concentrations were measured using enzyme-linked immunosorbent assay ELISA kit, DuoSet ELISA, R&D Systems, run on SDS-PAGE and validated using MALDI-TOF MS.

2.2.2 hCG Enrichment Methods

Three methodologies were investigated for the purification/enrichment of hCG collected from urine and conditioned media. The employed methods bind hCG according to their carbohydrate-binding properties (lectin) or by using solid phase-conjugated antibodies (protein G sepharose and paramagnetic beads).

2.2.2.1 Enrichment with lectin (Concanavalin A)

hCG enrichment was carried out using the Concanavalin A (Con A), Medicago AB, Uppsala, Sweden. This comprised of the Concanavalin A Lectin (*Canavalia ensiformis*) (MW=104kDa) immobilized onto agarose beads, which were isolated from Jack beans and purified by affinity chromatography. The ConA Lectin Resin was supplied as a white lyophilized powder with Binding/Wash buffer (0.1mM NaCl, 0.5 $MgCl_2$ 0.5 mM $MnCl_2$ and 0.5mM $CaCl_2$), no preservatives were added. 2mL of deionized water was added and the bottle was gently swirled until the resin was fully dissolved. 50μL of homogeneous suspension of ConA Lectin Resin was transferred to the tube and centrifuged (1 min, 1000 × *g*) following removal of the supernatant.

Next, 20 μL of JEG-3 conditioned media or pregnant urine samples were added to 50μL of the prepared ConA Lectin Resin along with 100μL Binding/Wash Buffer and samples were incubated at room temperature for 45 minutes on a slowly rotating platform. The mixture was then centrifuged (1 min, 1000 × g) and the supernatant was discarded, a further 100μL of Binding/Wash buffer was added. After thorough mixing (RT, 10min), once more the supernatant was discarded and 20μL of Elution Buffer was added to the resin mixed and centrifuged for a further 5 minutes. The resulting supernatant (1:1) was mixed with 15μl 2x Laemmli SDS-PAGE buffer, boiled for 10 min and loaded onto the (4-15%) SDS-PAGE gel for further analysis as described in Section 2.2.3.1.

2.2.2.2 Enrichment with anti-hCG conjugated protein G Sepharose

hCG enrichment was carried out using Protein G Sepharose (CL-4B), GE Healthcare conjugated anti-hCG beta antibodies 5004-SP-1, and 5008-SP-5, from Medix Biochemica (MedixMAB™, Finland). To check the optimal binding capacity of anti-hCG beta antibodies to Protein G Sepharose beads, 120μg of Protein G Sepharose powder was reconstituted with 10 mL of dH_2O, mixed gently and left to settle for 15 min. The supernatant was then discarded. This procedure was repeated three times with a final volume of 600 mL dH_2O added to make a heterogeneous mixture. Next, 500μL of the above suspension was transferred to an Eppendorf tube and centrifuged (500 x *g*, 3min). The excess supernatant was replaced with 50mM Tris

buffer (300 μL, pH 7.0), vortexed (1min) and discarded again. This wash process was repeated five times. Next, 5 x 50μL aliquots of this suspension were added to Eppendorf tubes containing 100μL of anti hCG antibody mix: anti-hCG beta 5004 SP-1 and anti-hCG beta 5008 SP-5. Following the preparation of various doses ranging from 30-500 g/mL, the starting concentration of the antibody mix was 1mg/mL (see Table 2.1), and samples were incubated overnight at 4°C. Supernatants (10μL) after centrifugation at 500 x *g* for 3 minutes were collected and mixed with 10μl 2x Laemmli SDS-PAGE buffer, boiled for 10 min and subjected to SDS-PAGE gel electrophoresis (Section 2.2.3.1).

Table 2.1 Optimisation of anti-hCG conjugation to protein G sepharose beads prior to immunopurification of hCG. Optimisation of antibody 5004-SP-1, and 5008-SP-5, from Medix Biochemica (MedixMAB™, Finland) concentration at 30-500 μg/mL range for binding to a constant amount of reconstituted protein G sepharose beads (50μL).

TUBE	Reconstituted beads (μL)	Stock of mix antibodies (1mg/mL) Volume (μL)	Volume of diluent (μL)	Final concentration of antibody mix (μg/mL)
1	50	75	25	500
2	50	37.5	62.5	250
3	50	18.75	81.25	125
4	50	9.37	90.63	62.5
5	50	4.68	95.32	31.2

Following the above optimisation experiment of antibody to beads ratio (Table 2.1), the next step was to establish antigen (hCG) binding capacity to the beads with constant antibody concentration and constant beads volume (50μL) as presented in Table 2.2. Therefore, 100μL of hCG standard (AC048) at various concentrations (0.3-4μg/mL) was added to the prepared overnight immunocaptured beads with anti-hCG beta antibodies. After overnight incubation at 4°C, supernatant from each sample was collected for further analysis and the remaining suspensions were washed three times with Tris buffer (50mM, 300μL, pH 7.0) and centrifuged (500x *g* for 3min). In order to elute hCG, the supernatant was removed and the sepharose beads were incubated for 30 min in RT with acetic acid (3M, 100μL) and neutralised with ammonium

hydroxide (1M, 290µL). Finally, 10µL of each sample (elution fractions) was mixed (1:1) with 2x Laemmli SDS-PAGE buffer, boiled for 10 min and loaded onto the gel.

Table 2.2 Binding capacity of Protein G Sepharose (CL-4B), *GE Healthcare* beads with attached conjugated anti-hCG beta antibodies (5004-SP-1, and 5008-SP-5, from Medix Biochemica (MedixMAB™, Finland) to hCG glycoprotein at 0.3-4 µg/mL concentration range.

TUBE	Volume reconstituted Ab coupled beads (µL)	Volume (µl) of hCG (20µg/mL)	Volume of diluent (µl)	hCG concentration (µg/mL)
1	50	30	20	4
2	50	15	35	2
3	50	7.5	42.5	1
4	50	3.75	46.75	0.5
5	50	1.87	48.13	0.3

2.2.2.3 Enrichment with hCG conjugated magnetic Dynabeads®

Enrichment of hCG was carried out using magnetic Tosylactivated Dynabeads® M-280 (14203, Invitrogen), with anti-hCG beta antibodies (5004-SP-1and/or 5008-SP-5 antibodies from Medix Biochemica (MedixMAB™, Finland).

The 165µL of paramagnetic tosyl activated (solid-phase) Dynabeads were washed three times in 0.1M sodium phosphate buffer, pH 7.4 (1mL). For each wash, after the addition of washing buffer, the tube was placed on the magnetic rack for 2 min and the supernatant was carefully removed. Washed beads were resuspended in 0.1M sodium phosphate buffer containing 100 µg anti-hCG beta antibodies of either 95µL (1.052 mg/mL) 5004-SP-1 or 19.6µL (5.1 mg/mL) 5008-SP-5 in a total reaction volume of 150µL. To achieve a concentration of 5mg beads/mL suspension (250µL), following manufacturer recommendation, 100µL of ammonium sulphate (3M, dissolved in 0.1M sodium phosphate buffer) was added. The mixture of antibody-conjugated beads (5004 and 5008 separately) was then incubated for 12, 24, 48 or 64 hours at room temperature on a rotating platform. Next, the supernatant was removed and beads were

washed to remove any unbound antibody. Following, the addition of 1mL PBS with 0.05% Tween 20 and second incubation at 37°C, 800 rpm for 2h. The supernatant was then removed, and beads were washed twice with 1ml of PBS-T (PBS containing 0.01% Tween 20) with vortex mixing between each wash. Finally, the beads-antibody complex was resuspended in PBS-T to a final volume of 250μL.

Following the antibody conjugation, as described above, 50μL of the antibody-coupled beads (1mg) were incubated with 0.05-4μg of hCG protein standard (AC048, Fitzgerald) (1mL, 50mM AMBIC, pH 8, RT) on a roller for 0, 2, 4, 8 and 24 hours to measure the percentage of hCG bound to the beads and subsequently to measure hCG recovery from the beads (Section 2.3.2.3 and 2.4). At each time point, protein concentration in the liquid phase (supernatant) was measured by removing 4μL of the sample exposed to the magnet. The protein concentration in the solution was quantified by NanoDrop™ Lite Spectrophotometer (Thermofisher, UK) or hCG beta DuoSet ELISA, DY9034-05 (R&D Systems, UK) for each collected sample.

Lastly, to elute hCG (0.005-4μg/mL) attached to Dynabeads, beads were washed following four steps of the washing process with 500μL of PBS with 0.05% Tween 20, 500μL of PBS x 2 and 500μL of 50mM AMBIC, vortex mixing and removing the supernatant between each wash. Next, hCG was eluted by the addition of TFA (15μL, 0.1%TFA, pH 2.5) following a continuous shake (1000rpm) at room temperature for 10 minutes. Then, the solution was neutralised using 13mM ammonia solution (33.6μL, pH 10.3) and 50mM ammonium bicarbonate (20μL, pH 8.4) to obtain a mixture of beads and detached protein (128.6μL, pH 9.4). Samples were placed on the magnet (approximately 2 min), and the supernatant was transferred to a new collection tube and left to dry in the vacuum concentrator (1.5h, 30°C). Samples were then stored at -20°C. For further analysis, such as SDS-PAGE, ELISA or MALDI-TOF MS eluted proteins were reconstituted in ultrapure water following methods described in Section 2.2.3.

2.2.3 Detection of hCG

For hCG detection in crude samples such as cultured media and human urine (before and after enrichment), several analytical methods such as gel electrophoresis, enzyme-linked immunosorbent assay and mass spectrometry were applied as described in the sections below.

2.2.3.1 Sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE gel electrophoresis is a simple and rapid analytical tool that can separate negatively charged protein molecules or peptides transported through a polyacrylamide gel by an electrical field accordingly to their size. Proteins carry a net charge at any pH other than their isoelectric point and will migrate at a rate proportional to their charge density. To perform SDS-PAGE gel electrophoresis, proteins were either loaded directly on the gel, or proteins' peptide chains were firstly cut at the carboxyl side of the arginine and/or lysine residues using trypsin following the addition of 2-mercaptoethanol (10% w/v) which allowed breaking of disulphide bonds and completing saturation of the polypeptide. Thus, reduced and negatively charged peptides travel through an electric field according to their differences in molecular weight and their charge to mass ratio. Subsequently, the two layers in the polyacrylamide gel served different purposes: the stacking gel was used to concentrate/pack all of the proteins/peptides into a single band, allowing them to all start migration at the same time, and resolving gel separated proteins/peptides based on their molecular weight. The distance moved on the gel was proportional to $\log_{10}$ molecular weight under the used conditions. By comparing the distance travelled through the gel, with the prestained marker of known molecular weight (ladder) run on the gel (Precision Plus Protein Standards Dual Colour Ladder, Bio-Rad, UK), the approximate size fragments of glycoprotein separated by SDS-PAGE were determined.

To detect hCG and other proteins and/or its peptides using SDS-PAGE gel electrophoresis three approaches were used. Firstly, to detect proteins, commercially available hCG protein standards and/or crude samples such as cultured media and human urine were mixed (1:1) with 2x Laemmli SDS-Page buffer, boiled for 10 min and 20 µL of each sample was loaded onto the (4-15%) SDS-PAGE gel (120V, 40min). Secondly, to detect peptides, the tryptic digestion procedure of hCG (1µg, 50mM AMBIC, pH 8) and other proteins in the solution was undertaken. Containing denaturation of protein (boil, 10min), addition of 10mM DTT (10µL, 56°C, 10min), 20mM IAA (10µL, RT, 10min) and trypsin 20:1 ratio (20µg/mL in 50mM AMBIC with 2% ACN, 37°C, 19h) accordingly, following incubation at 37°C, overnight.

Trypsinised samples were mixed (1:1) with 2x Laemmli SDS-Page buffer, and boiled for 1 min. Next, 20 µL of each sample was loaded onto the SDS-PAGE gel (120V, 40min). After initial gel electrophoresis of crude samples (urine, conditioned media) and/or protein standards, the bands corresponding to proteins' molecular weight of interest i.e. $\geq 15kDa \leq 50kDa$ were cut out from the gel and further processed. Gel fragments were immediately saturated in 100µL of 50% acetonitrile with 100mM ammonium bicarbonate (AMBIC) solution for 10min with sonication, washed with water and dehydrated (100% ACN, 5min x 3, sonication). After air-drying (20min), reduction (10mM DTT in 100mM AMBIC, 56°C, 30 min) alkylation (55mM IAA in 100mM AMBIC, RT, 45 min in dark) and final dehydration (100% ACN, 15min x 2, air-dry) took place. Next gel sections were digested using trypsin (3.75µL, 20µg/mL) and incubated at 37°C overnight. The supernatant was collected, mixed with organic matrix (1:1) and spotted on the MALDI-TOF target plate (MTP 384 ground steel) and air-dried as described in Section 2.2.3.3.

2.2.3.2 Enzyme-linked immunosorbent assay (ELISA)

To quantify hCG concentration in conditioned media and pregnancy urine samples hCG ELISA Kit from Calbiotech. Inc., USA and the human hCG beta DuoSet ELISA, DY9034-05, R&D Systems, UK were used. Both assays were run based on solid-phase sandwich ELISA as a recommended protocol by the manufacturer to quantify intact hCG and hCGβ subunits. The biotin labelled anti-hCG and the anti-hCG hRP conjugate was added to the wells coated with Streptavidin. The hCG bound to the anti-hCG antibodies formed "a sandwich" on the streptavidin-coated plate. Unbound proteins and HRP conjugate were washed off by wash buffer, upon the addition of substrate.

Firstly, the assay procedure for hCG ELISA Kit from Calbiotech. Inc., USA microplate coating at room temperature for 30 min before use. 25µL of samples, hCG standards and controls were added to the individual wells, following 100µL conjugate reagent and 60 minutes incubation at RT. Next, the plate was washed three times with 300µL of 1x wash buffer and blotted on an absorbent paper towel. After the addition of 100µL TMB substrate, the plate was incubated for 15 min at RT and stop solution (50µL) was added to each well for 15 mins. Absorbance was read at 450nm standard curve range was between 10-250mIU/mL.

Secondly, the assay procedure for human hCG beta DuoSet ELISA, DY9034-05, Lot number P247055, R&D Systems, UK, comprised of multiple reagents at a specific working concentration, i.e. wash buffer (0.05% Tween20 in PBS), detection antibody reconstituted in reagent diluent (500ng/mL), standard hCG reconstituted in reagent diluent (7.81-2000 pg/mL),

streptavidin-HRP diluted in reagent diluent (1:40) and capture antibody reconstituted in PBS (120µg/mL) for coating a microplate (100µL per well, RT, overnight) all based on the manufacture protocol. After washing a 96 well microtitre plate (Maxisorp, Thermofisher, UK) three times with 300µL wash buffer, reagent diluent (300µL) was added and incubated for 2 hours at room temperature to reduce non-specific binding, next it was washed three times more. 100µL of biological samples (diluted with reagent diluent at 1:1000, 1:10 000 and 1:100 000), hCG standards and controls were added to the individual wells, following 2 hours incubation and wash (x3). Next 100µL of the detection antibody (2 hours, RT), and 100 µL of streptavidin-HRP (20 min, RT) were added, followed by washing a plate twice with wash buffer before each step. Finally, 100 µL of substrate solution (20min, RT) and 50µL of stop solution (1N sulphuric acid) were added and the absorbance was read at 450nm within 15min.

2.2.3.3 Matrix-Assisted Laser Desorption/Ionisation Time of Flight (MALDI TOF) Mass Spectrometry (MS)

Matrix-Assisted Laser Desorption/Ionisation (MALDI) Time of Flight (TOF) Mass Spectrometry (MS) was used to identify hCG and any other proteins by matching their peptide fragment masses to the theoretical peptide masses generated from protein databases (peptide mass fingerprinting) as presented in Chart 2.1.
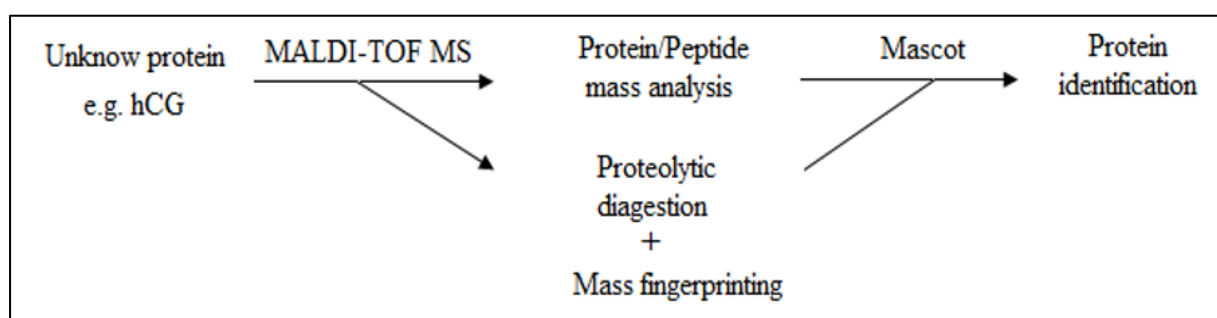
Chart 2.1 Diagram shows protein/peptide identification by MALDI-TOF-MS. Samples were mixed with 10mg/mL SA or 10mg/mL HCCA: DHB matrix for analysis of proteins and the tryptic peptides, respectively. To identify a single protein, the Mascot software search engine and SwissProt database containing amino acid sequences of known proteins were used.

The first step of sample analysis with the MALDI-TOF mass spectrometry was desorption and ionisation. The analyte was mixed in an organic compound matrix spotted on a solid surface

made of a conducting metal (target plate). When a direct laser UV light (nitrogen laser – 337nm) was shot onto the analyte, the irradiated spot was rapidly heated and become vibrationally excited. Absorption of light by the matrix (ablation process), transfers energy (protons) to the analyte and ionized sample, triggering the formation of a dense cloud mixture (matrix/analyte) into the gas phase. Next, the analyte could be either protonated or deprotonated by direct collisions between excited ions in the matrix. This soft ionization technique, in which protons travel from the target plate towards the detector with a fixed potential difference (20-30 kV) and the same initial kinetic energy carries the most common single or double positive charge $[M+H]^+$ or $[M+2H]^{2+}$. Therefore, ions were separated accordingly to their velocity along the time of the flight tube (TOF) and their proportional mass to charge ratio (m/z). The m/z=$2Ue$ (t/l) 2, where (U, e and l are constant), which means the ion m/z is proportional to the flight time. [U is the overall voltage, e is an elemental charge (1.602 x10-19 Coulomb), t is time and l is the length of the tube]. Finally, ions were identified using a detector (Figure 2.1) (Gross, 2006, de Hoffmann and Stroobant, 2007; Kicman *et al.,* 2007; Gosrani, 2014).

Therefore, in the following experiments, peptide mass fingerprinting (PMF) using Matrix-Assisted Laser Desorption/Ionisation (MALDI) TOF MS/MS for hCG identification was applied. 1µL of the sample of hCG standard material or enrichment from urine or conditioned medium was mixed with (1:1) sinapinic acid (SA) or alpha-cyano-4-hydroxycinnamic acid (HCCA) matrix with 2,5-Dihydroxybenzoic acid (DHB) and spotted onto a 384 ground steel MALDI plate. Protein and peptides identification settings were as follows: the linear detector gain (63 x 3130V), positive mode with m/z rage of 700 Da-60 kDa, 6_ultra smartbeam parameter set with the zoom range of ±5% and reflector detector gain (29 x 2113V), positive mode with m/z range of 700-5000Da, 6_ultra smartbeam parameter set with the zoom range of ±5%. Peptide sequences were analysed using Mascot search engine searches. Mascot search parameters included SwissProt database, NCBI BLAST taxonomy browser limited to *Homo sapiens* species. Protein view: CGB7_Human, Choriogonadotrop subunit beta 7 OS=Homo sapiens OX=9606 GN=CGB7 PE=2SV=1. Experimental mass values were set to monoisotopic and variable modifications such as carbamidomethyl (C) and oxidation (M). The error window on the experimental peptide mass value was set to 500ppm. Mass values corresponded to charge carrier $[M+H]^+$. The number of missed trypsin digest cleavages was set to 1, which is optimal for founding the target protein against the experimental data.
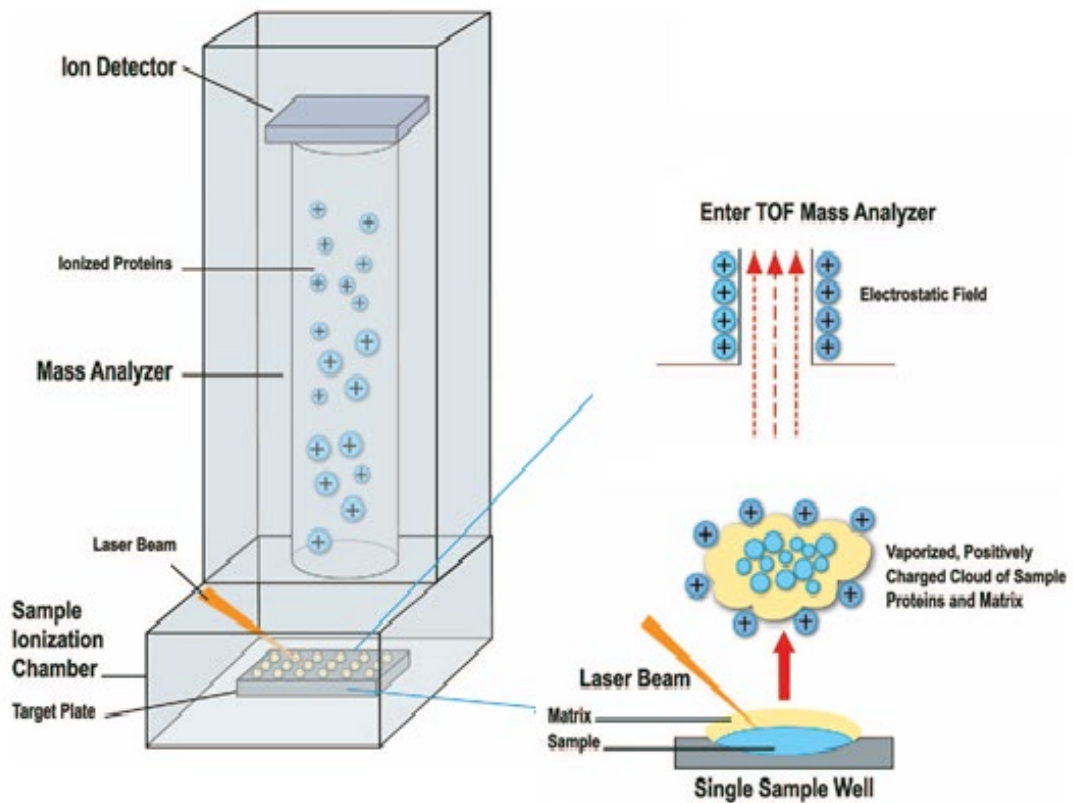
Figure 2.1 The overview of MALDI-TOF mass spectrometer containing ion detector, time of flight tube, laser bean and sample ionisation chamber. Processes of desorption and ionisation are shown on the target plate (on the right), where a positively charged cloud of the sample (proteins/glycans) and matrix travel to a mass analyser (Adapted from Patel *et al*., 2015).

2.3 Results

2.3.1 hCG detection prior to enrichment

In order to detect hCG in crude conditioned media and human urine samples, SDS-PAGE gel electrophoresis, hCG ELISA (Calbiotech, USA) and MALDI-TOF techniques were employed for analysis.

2.3.1.1 hCG detected in conditioned media

hCG was detected in all cell lines tested and in all types of used media. Figure 2.2 shows a comparison of hCG levels secretion between the JEG-3 choriocarcinoma cell line and the extravillous trophoblast HTR-8 used as a control. hCG levels secreted by JEG-3 were significantly higher than the HTR-8 cell line, (JEG-3 11.73-58.58 mIU/mL; HTR-8 0-2.9 mIU/mL) into MEM serum-free medium over 42 day period with 20% supplementation of fresh media.
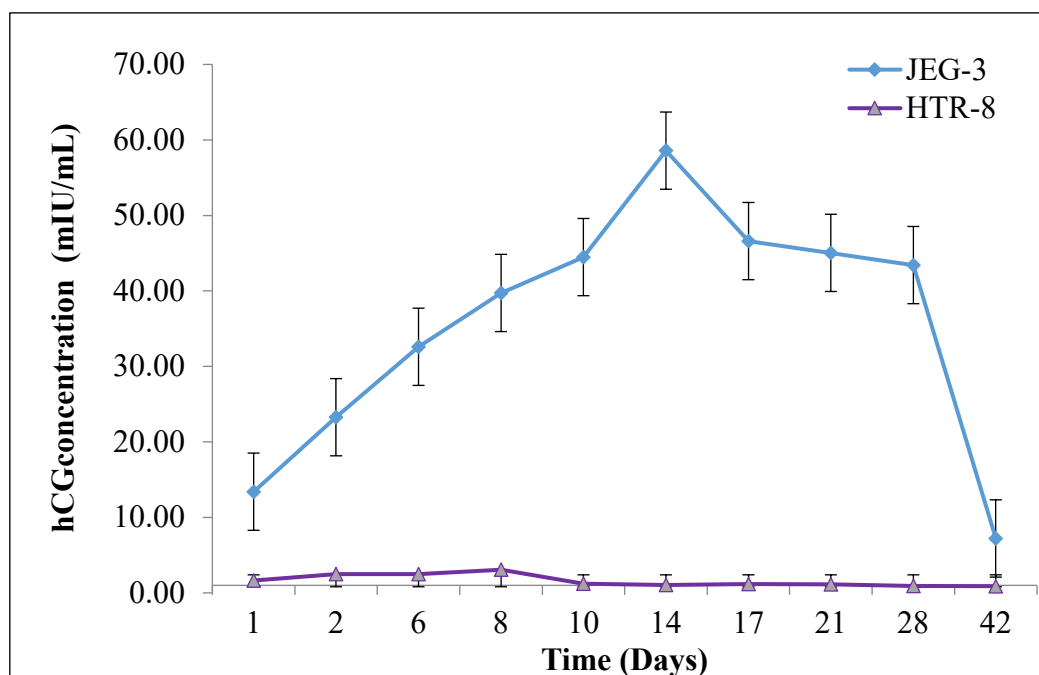

Figure 2.2 hCG concentration (mIU/mL) secreted into MEM serum-free medium by JEG-3 and HTR-8 cell lines over 42 days period, measured using hCG ELISA (Calbiotech. Inc., USA). The mean concentration of hCG in JEG-3 was 35.79 mIU/mL, SD=14.72. For HTR-8, the mean was 1.92 mIU/mL, SD=0.64. SD-standard deviation, error bars represent the standard error of the mean (n=10) (SEM)

Effect of cells density on hCG secretion into a medium showed significant differences in hCG concentration over time. Figure 2.3 illustrates the mean hCG concentrations secreted into MEM serum-free medium by 12×10^6 (blue line) and 8×10^6 (orange line) cells over 42 days in a $175 cm^2$ flask. There was a steady increase in hCG detected over the first 12-14 days. Under the used conditions, the hCG concentration (mIU/mL) was significantly higher at the lower cell density with mean hCG levels at day 14 of 250 mIU/mL and 75 mIU/mL respectively, $p < 0.001$. After day 14, hCG yields decreased at both cell densities.
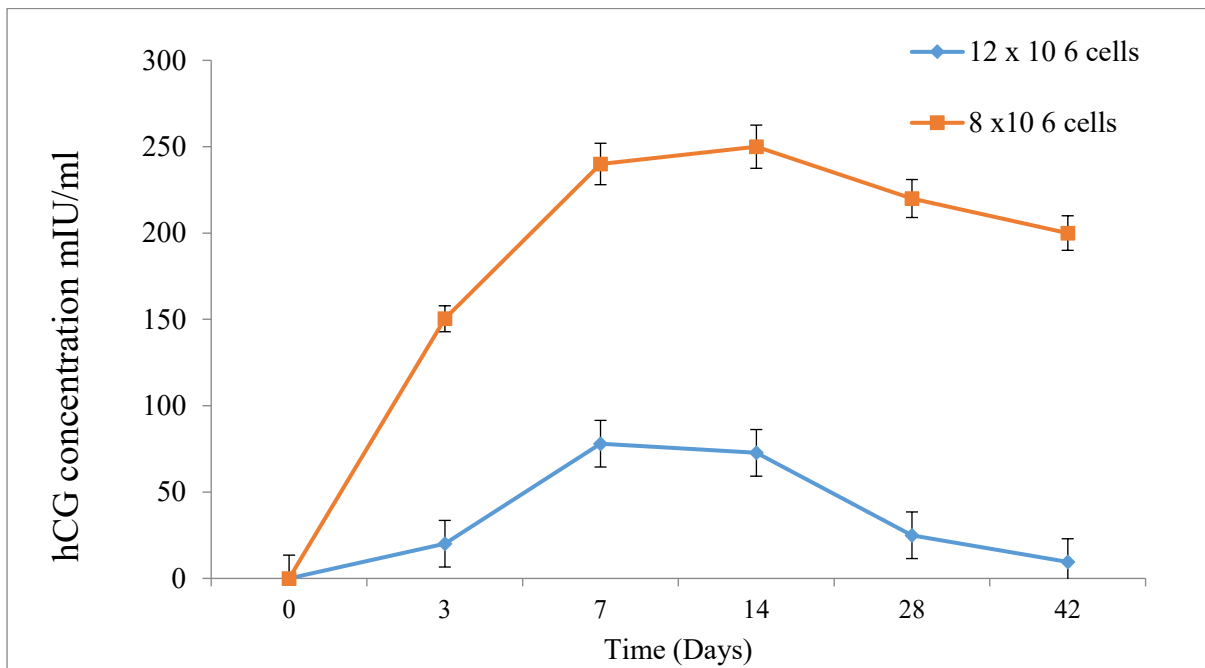

Figure 2.3 hCG concentration (mIU/mL) secreted into MEM serum-free medium by JEG-3 at 8×10^6 cells and 12×10^6 cell density, over 42 days period ($175 cm^2$ flasks), measured using hCG ELISA (Calbiotech. Inc., USA). Error bars represent the standard error of the mean (SEM), n=6, $p < 0.001$

Thus, to collect the maximum hCG yield at the optimal cell density, JEG-3 (8×10^6) cells were passaged into 175cm^2 and grown for 14 days with a MEM-SF medium. Additionally, to enhance hCG yield three different conditioned media including ITS medium, MEM serum-free medium and MEM medium containing 10% FCS were tested as shown in Figure 2.4. Again the peak of hCG concentration was established at 14 days mark for all three media. The mean hCG concentration detected in MEM-SF and ITS were 148.4 ($\pm$ 1.4) and 97.4 ($\pm$ 7.2) mIU/mL respectively, and 41.1 ($\pm$ 3.7) mIU/mL for MEM-10% FCS. Moreover, in this data sample, consistent, overtime hCG secretion into conditioned media was observed with an average hCG concentration of 47.5 mIU/mL, 130.3 mIU/mL and 291.2 mIU/mL on the 3rd, 7th and 14th day of the experiment. The association of hCG secretion (on day 14) from each conditioned medium revealed a statistically significant difference ($p < 0.008$).

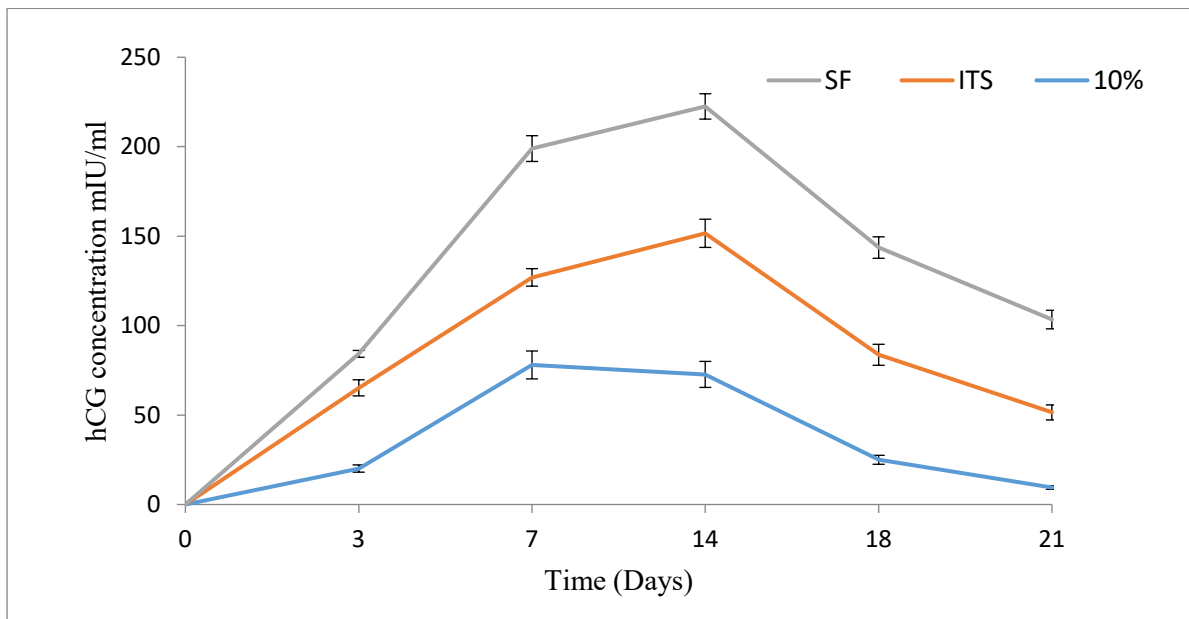**Figure 2.4** Measurement of hCG concentration secreted by JEG-3 in various media: 10% FCS, ITS and SF (8 x 10^6 cells in 175 cm^2 flasks) measured using ELISA (Calbiotech. Inc., USA). Error bars represent the standard error of the mean (SEM), n=5

Subsequently, hCG isoforms secreted from JEG-3 choriocarcinoma cells and other proteins commonly detected in the conditioned media were separated using sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) and presented in Figure 2.5. Multiple bands were observed on the gel from the JEG-3 cell line cultured in ITS medium, MEM serum-free medium and MEM medium containing 10% FCS, collected on day 14 (Figure 2.5, Lane 2, 3 and 4). On lane 5 (Figure 2.5) the standard sample (Section 2.2.1.1) of 1ug intact hCG (AC048) standard was loaded. The three bands at ≤50 kDa corresponding to intact hCG and α and β subunits were observed. The accurate molecular weight of each band was determined by the equation used for measuring the relative mobility of protein bands in the gel, in this case, i.e. $y = -0.1244x + 5.6181$, $R^2 = 0.993$, which established subunits: α at 15 kDa and β at 25.7 kDa and intact hCG at around 37.4 kDa.
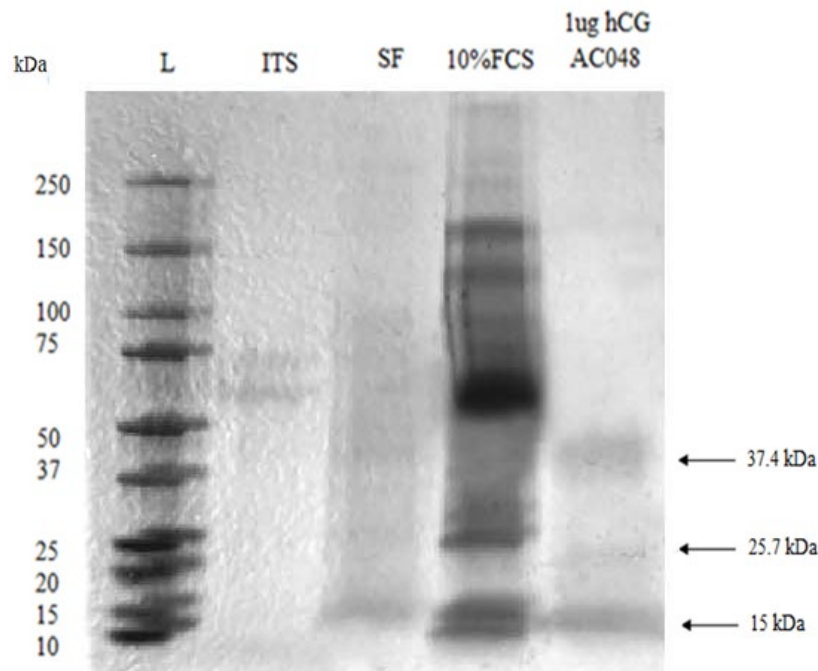
Figure 2.5 SDS-PAGE gel of three conditioned media from JEG-3 choriocarcinoma cells (8×10^6) collected on day 14. Lane 1: L-ladder, Lane 2: ITS-Insulin, Transferrin and Selenium medium supplement, Lane 3: SF-MEM serum-free Dulbecco's Modified Eagle's Medium, Lane 4: 10%FCS- Dulbecco's Modified Eagle's Medium containing 10% fetal calf serum, Lane 5: 1µg hCG standard (AC048). Three bands labelled with arrows at 15, 25.7 and 37.4 kDa correspond to α and β subunit, and intact (α+β) hCG respectively.

Next, the sections of a gel corresponding to hCG molecular weight defined earlier i.e. below the 50 kDa mark and above the 15 kDa mark were cut out from the gel (Figure 2.5) and were analysed by the peptide mass fingerprinting (PMF) technique using MALDI- TOF described in Section 2.2.3.3. In Figure 2.6 the presence of hCG was confirmed by distinctive peaks visible on the chromatograms of around 15-18 kDa corresponding to α subunit, 22-23 kDa corresponding to β subunit, and 36-46 kDa peak of intact glycosylated hCG. Also in conditioned media samples, the double-charged peak of around 33.3 kDa $[M+2H]^{2+}$ from albumin was detected by MALDI-TOF. This might be due to sample contamination with albumin while cutting the section from a gel.
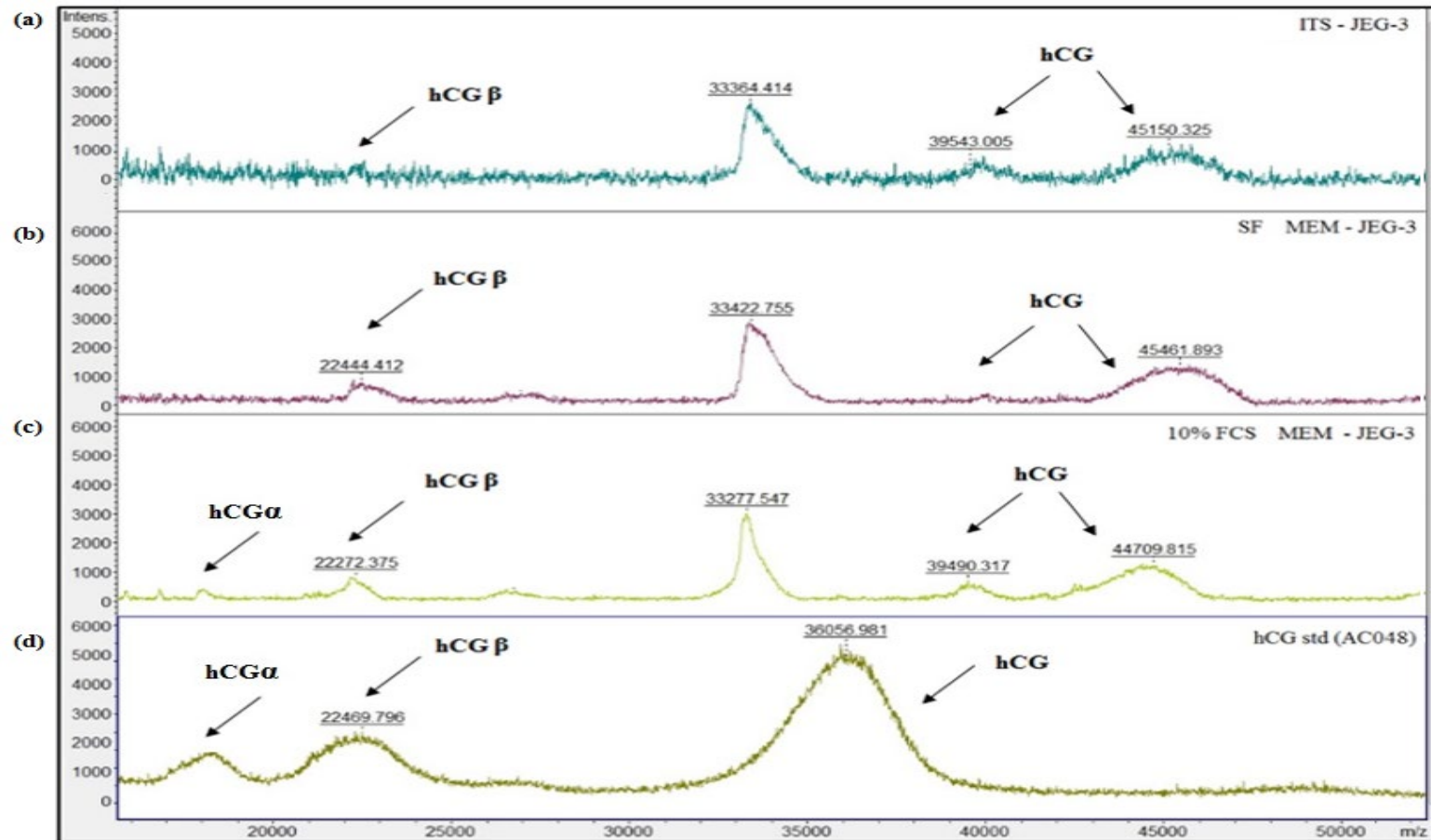
Figure 2.6 MALDI-TOF MS mass spectra of cut out bands (≥15 kDa ≤ 50kDa) from SDS-PAGE gel of three conditioned media from JEG-3 choriocarcinoma cells (8×10^6) collected on day 14. (**a**) ITS-Insulin, Transferrin and Selenium medium supplement, (**b**) SF-MEM serum-free Dulbecco's Modified Eagle's Medium, (**c**) 10%FCS- Dulbecco's Modified Eagle's Medium containing 10% fetal calf serum, (**d**) 1μg/mL hCG standard (AC048). Peaks were corresponding to α subunit (15-18 kDa), β subunit (22-23kDa) and intact hCG (36-46kDa). The peak at 33kDa $[2H]^{2+}$ matches the molecular weight of albumin commonly found in conditioned media samples.

The gel section of bands with molecular weights (≥15 kDa ≤ 50kDa) from the SDS-PAGE gel shown in Figure 2.5 were further trypsinised following the method in Section 2.2.3.1 and the presence of hCG matching peptides was confirmed by the PMF analysis in each sample (Table 2.3). MALDI-TOF MS/MS analysis of hCG secreted by JEG-3 choriocarcinoma cells detected specific hCG peptides matching theoretical peptide peaks list derived from the SwissProt database in silico digestion. The best match of 89% hCG amino acid sequences similarity was determined by the Mascot software search engine and NRBD protein database from NCBI.

Table 2.3 hCG peptide mass fingerprinting identification of extracted peptides from SDS-PAGE gel using MALDI-TOF MS/MS **(a)** The full sequence of *Choriogonadotropin subunit beta variant, CGB3, Homo sapiens*, found in ITS, MEM SF and MEM 10%FCS JEG-3's conditioned media after 14 days incubation, subjected to SDS-PAGE, cut out from the gel and digested with trypsin as in Section 2.2.2.1. The amino-acid sequence in red indicates the similarity of the published database beta hCG subunit amino-acid sequence (89%). **(b)** Table of eleven detected peptides matching hCG amino-acid sequence detected from JEG-3 conditioned medium.

(a)

1	MEMFQGLLLL LLLSMGGTWA SREMLRPRCR PINATLAVEK EGCPVCITVN
51	TTICAGYCPT MTRVLQGVLP ALPQVVCNYR DVRFESIRLP GCPRGVNPVV
101	SYAVALSCQC ALCRRSTTDC GGPKDHPLTC DDPRFQASSS SKAPPPSLPS
151	PSRLPGPSDT PILPQ

(b)

Aminoacid Position number (Start-End)	Molecular weight	AA Sequence of Peptides + (modifications) (C)- Carbamidomethylation, (M)- Oxidation
1-22	2467.02	-.MEMFQGLLLLLLLSMGGTWASR.E
11-49	3832.77	R.LLLSMGGTWA SREMLRPRCR PINATLAVEK.E +(M)
23-28	816.43	R.EMLRPR.C + (M)
41-63	2678.95	K.EGCPVCITVNTTICAGYCPTMTR.V+ (M); 3x(C)
64-80	1924.92	R.VLQGVLPALPQVVCNYR.D
84-88	650.39	R.FESIR.L
89-94	698.38	R.LPGCPR.G
89-114	2904.14	R.LPGCPRGVNPVVSYAVALSCQCALCR.R + (C)
95-114	2223.88	R.GVNPVVSYAVALSCQCALCR.R + (C)
115-134	2283.95	R.RSTTDCGGPKDHPLTCDDPR.F
135-142	840.02	R.FQASSSSK.A

2.3.1.2 hCG detection in urine from healthy pregnancy and choriocarcinoma

After detection and measurement of hCG secreted into JEG-3 conditioned medium using SDS-PAGE, hCG ELISA and MALDI-TOF, hCG levels in human urine from healthy pregnancy and choriocarcinoma patients using identical methods prior to enrichment were applied and are presented below.

In Figure 2.7 multiple bands were detected on an SDS-PAGE gel i.e. Lane 2: 1st-trimester pregnancy urine, Lane 3: urine from choriocarcinoma and Lane 4: 1μg of intact hCG (AC048) standard. The bands corresponding to intact hCG (37.4 Da), β (25.7 Da) and α (15 Da) subunits were observed however, additional bands in urine samples were difficult to distinguish. Further analysis by MALDI-TOF confirmed the presence of albumin, creatinine and uromodulin in urine samples. Average hCG concentrations, as measured by ELISA, from 5 participants in first-trimester pregnancy urine compared to the second and third trimester were (mean±SEM) 11800.9 ± 4026 mIU/mL, 9054.0 ± 4106.0 mIU/mL and 21313.8 ± 4343.3 mIU/mL, respectively. Thus, the variation between the sample from multiple participants at various weeks was not significantly different (ANOVA, $p=0.12$). Additionally, the average hCG concentration in choriocarcinoma urine samples (5 participants = 9164.2 ± 912.6 mIU/mL) to healthy pregnancy was also not significantly different (ANOVA, $p=0.15$).
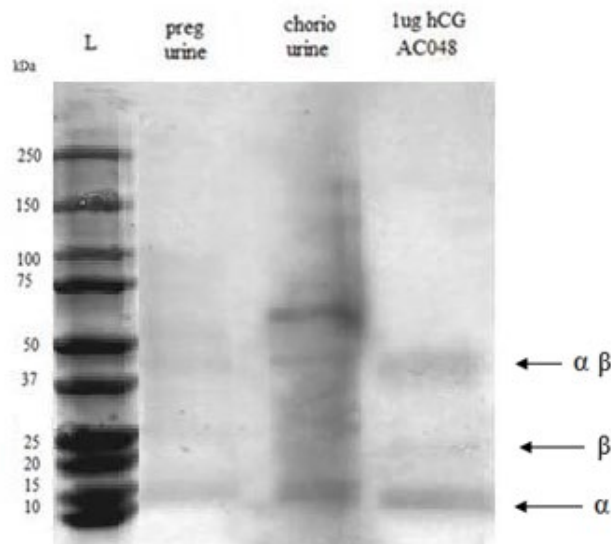

Figure 2.7 SDS-PAGE gel showing separation of crude urine from healthy pregnant female (Lane 2) and urine from choriocarcinoma patient (Lane 3). Lanes 1: ladder (10-250kDa), Lane 4: 1μg intact hCG (AC048) standard. The α and β subunit, and intact (α+β) hCG corresponds to 15kDa, 25.7 kDa and 37.4 kDa respectively.

2.3.2 Analysis of Enrichment Methods for hCG

2.3.2.1 Lectin Concanavalin A (ConA)

The serum-free (SF)- JEG-3 conditioned media, pregnant urine sample (1st trimester, week 7) and standard intact hCG (AC048) were enriched by incubation with Lectin Concanavalin A (ConA) (Section 2.2.2.1). Figure 2.8 presents an SDS-PAGE gel with respective three samples enriched by using a protocol with Concanavalin A, control samples of pure Concanavalin A only and hCG standard. Firstly in Figure 2.8, Lane 1 shows a band at around 100kDa corresponding to Concanavalin A molecular weight (Senear and Teller, 1981). Next, Lanes 3, 4 and 5 with three tested samples showed no detectable bands. Lane 7 correspond to the approximate molecular weight of intact hCG and α and β subunit. Further investigation of lanes 4, 5 and 6 involving cutting out the bands and MS fingerprinting failed to detect any peptide sequences corresponding to the presence of hCG (data not shown).
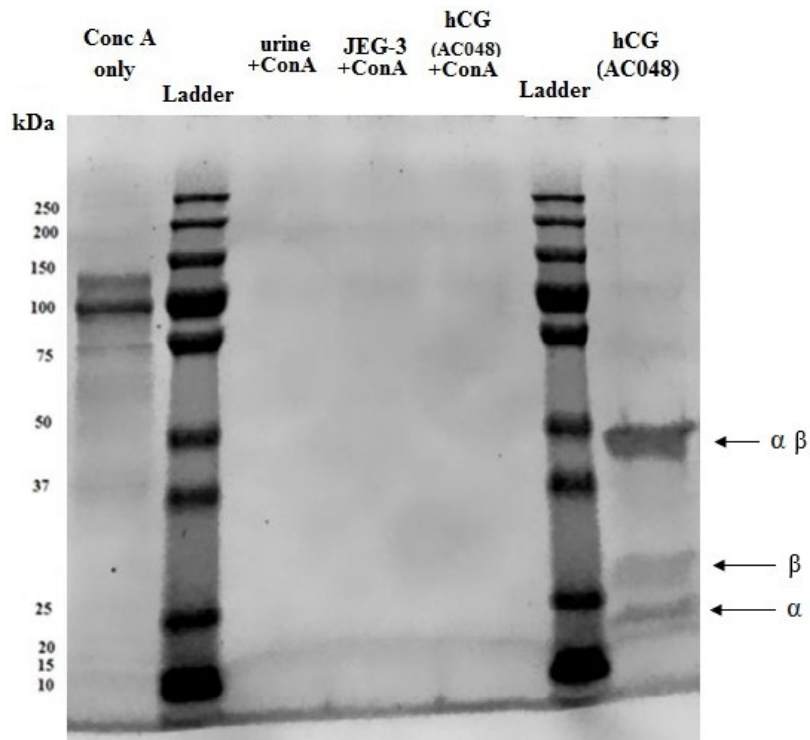

Figure 2.8 Enrichment of hCG using Concanavalin A (ConA) lectin. Lane 1: Concanavalin A (ConA) beads only, Lane 3: pregnancy urine enriched with ConA, Lane 4: JEG-3 enriched with ConA, Lane 5: hCG (AC048) standard enriched with ConA, Lane 7: 1µg of hCG standard (AC048), Lane 2 and 6: protein ladder (10-250kDa).

2.3.2.2 Protein G Sepharose magnetic beads

This method has comprised the optimisation of two stages. The first of these was to establish the capacity of the binding of the anti-hCG antibody to sepharose beads via analysis of the unbound fraction in supernatants. The supernatants removed from protein-G coupled sepharose beads after incubation with hCG Ab were analysed on SDS-PAGE gels. It was found that there was a significant amount of the antibodies had not bound when 50μL of beads were incubated with a 500μg/mL antibody (Figure 2.9, Lane 3). However, this amount was significantly less than the initial amount of Ab present (500 ug/ml, control Ab, Lane 2). When beads were incubated with lower concentrations of antibody (33.1-250 μg/mL) it was found that there were still significant amounts of unbound hCG antibodies left. To confirm the presence of the antibody on the SDS-PAGE gel, strongly visible bands at around 25 kDa and 50 kDa molecular weight were cut out, trypsinised following the method in Section 2.2.3.1 and spotted on the target plate for MALDI-TOF analysis (presented in Section 2.2.3.3). The presence of heavy and light chains was confirmed. In Table 2.4 the presence of immunoglobulin heavy constant gamma 1 (IgG) was presented following peptide sequence mass analysis using the Mascot search engine.
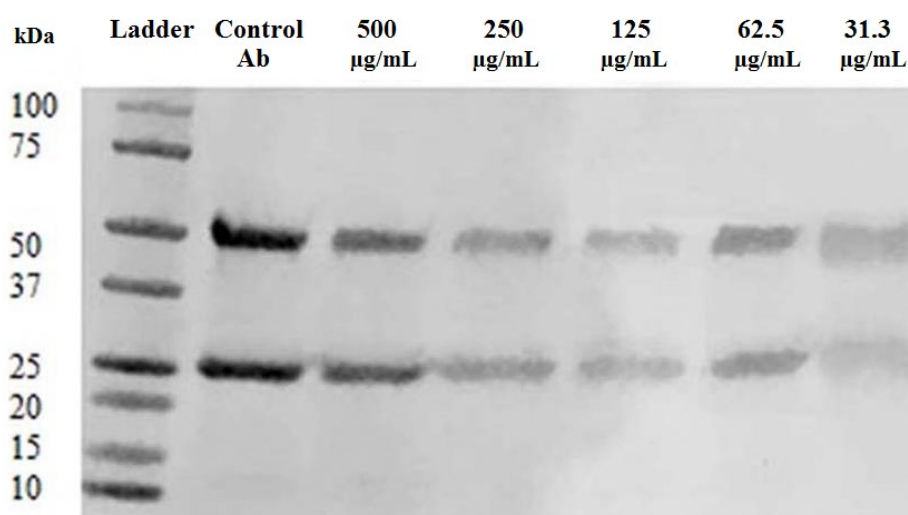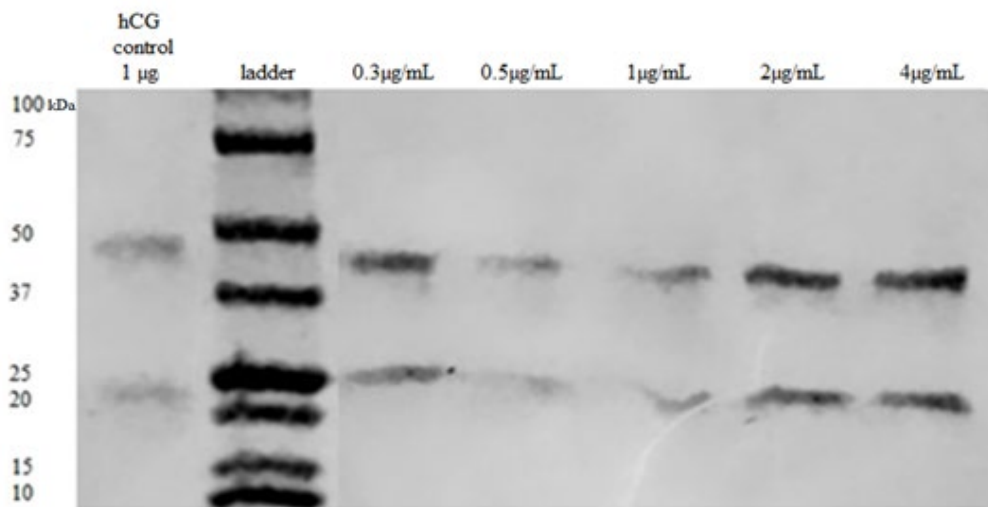

Figure 2.9 Unbound fractions collected from incubation of 50μL protein-G coupled sepharose with anti-hCG antibody (at the concentration range between 31.3 μg/mL to 500 μg/mL) and separated by SDS-PAGE with Ladder-protein marker (10-100kDa), Control antibody: 500μg/mL of anti-hCG antibody.

Table 2.4 Anti-hCG (IgG) antibody, peptide mass fingerprinting identification using MALDI-TOF MS/MS. The sequence of *Immunoglobulin heavy constant gamma 1 (IgG), Mouse,* chain was obtained after cutting bands at 50 kDa mark from the SDS-PAGE gel, subjected to trypsinisation as per Section 2.2.2.1 (Figure 2.9, Lane 2). The amino-acid sequence in red shows 95 % similarity to the published IgG amino-acid sequence protein database from NCBI.

1	ASTKGPSVFP LAPSSKSTSG GTAALGCLVK DYFPEPVTVS WNSGALTSGV
51	HTFPAVLQSS GLYSLSSVVT VPSSSLGTQT YICNVNHKPS NTKVDKKVEP
101	KSCDKTHTCP PCPAPELLGG PSVFLFPPKP KDTLMISRTP EVTCVVVDVS
151	HEDPEVKFNW YVDGVEVHNA KTKPREEQYN STYRVVSVLT VLHQDWLNGK
201	EYKCKVSNKA LPAPIEKTIS KAKGQPREPQ VYTLPPSRDE LTKNQVSLTC
251	LVKGFYPSDI AVEWESNGQP ENNYKTTPPV LDSDGSFFLY SKLTVDKSRW
301	QQGNVFSCSV MHEALHNHYT QKSLSLSPGK

Secondly, hCG binding to antibody-coupled sepharose beads was investigated using SDS-PAGE. Following overnight incubation, the liquid phase washed from each hCG sample (0.3 to 4 µg/ml) prior to elution was tested and any remaining unbound hCG was explored (Figure 2.10 a). Next, hCG was eluted from the beads by using 3M acetic acid and further neutralized with 1M ammonium hydroxide. As presented in Figure 2.10 (b) two bands at 20-25kDa and 37-50kDa were observed. The amount of hCG eluted from the beads was represented in the varied intensity of the bands which seemed to be strongest in the highest concentration of 4 µg/mL (but not quantified by densitometry). Eluted hCG was additionally confirmed by MALDI-TOF peptide mass fingerprinting and matching peptides were identified (data not shown).

(a)

(b)
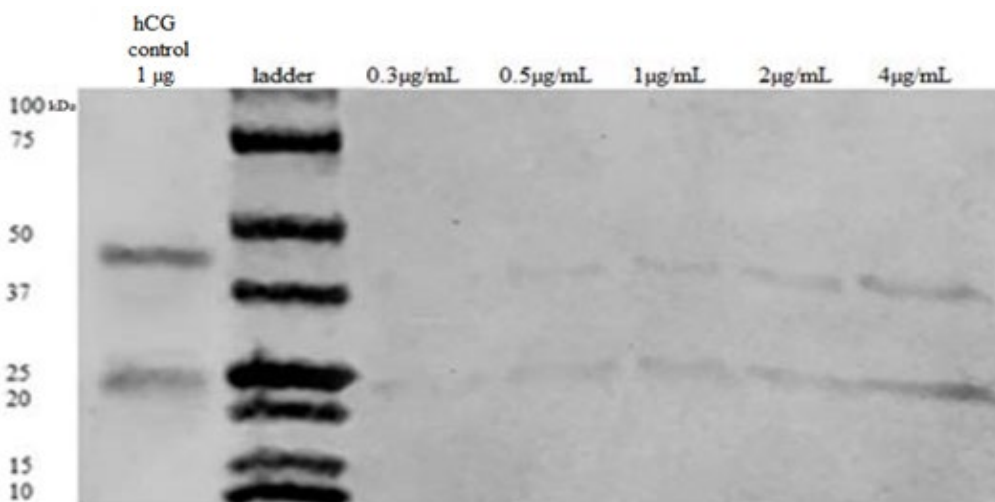
Figure 2.10 Results of hCG binding to antibody-coupled sepharose beads using SDS-PAGE gel **(a)** unbound fractions of hCG in the liquid phase (Lanes 3-7) after overnight incubation collected from samples containing 50µL of 125µg/mL antibody-coupled beads and hCG at concentrations range of 0.3-4µg/mL **(b)** eluted hCG (0.3 - 4 µg/mL) from the sepharose beads after 30 min incubation at room temperature with 3M acetic acid and neutralisation using 1M ammonium hydroxide (Lines 3-7). Lane 1: 1µg hCG standard (AC048), Lane 2: Ladder-protein marker (10-100kDa).

2.3.2.3 Magnetic Tosylactivated Dynabeads®

The effectiveness of conjugation of anti-hCG beta antibodies magnetic tosylactivated Dynabeads with a range of hCG concentrations (0.05-4µg/mL) was tested following the method described in Section 2.2.2.3. Similarly to the previous method, this was a two-staged process that required the linking of anti-hCG antibodies to tosylactivated paramagnetic Dynabeads (overnight, RT) followed by hCG binding to antibody-coupled Dynabeads (overnight, RT). Proteins eluted from the Dynabeads were assessed based on their molecular weight and results were presented using SDS-PAGE in Figure 2.11. Eluted protein fractions from various concentrations used were detected on SDS-PAGE as two bands with masses of intact hCG and β subunit. The bands were detectable from the lowest concentration used of 0.3 µg/ml of starting material (Figure 2.11). It was confirmed, following the PMF method described before in Section 2.2.3.3, that detected bands are matching a sequence of intact hCG (mass around 37 kDa) and hCG β subunit (≤ 25 kDa). Results are presented in Fig. 2.11. Lane 1 to 8 shows samples at 0.05 - 4µg/mL concentration of standard hCG with the constant amount of magnetic Dynabeads added in each incubation. The proportionally increasing intensity of the bands was detected for concentrations rang 0.3 - 4 µg/mL and no bands were detected below 0.3 µg/mL of used hCG. In conclusion, analysis confirmed that hCG bound successfully to antibody-coupled Dynabeads and was eluted with the lowest detection level of 0.3 µg/mL.
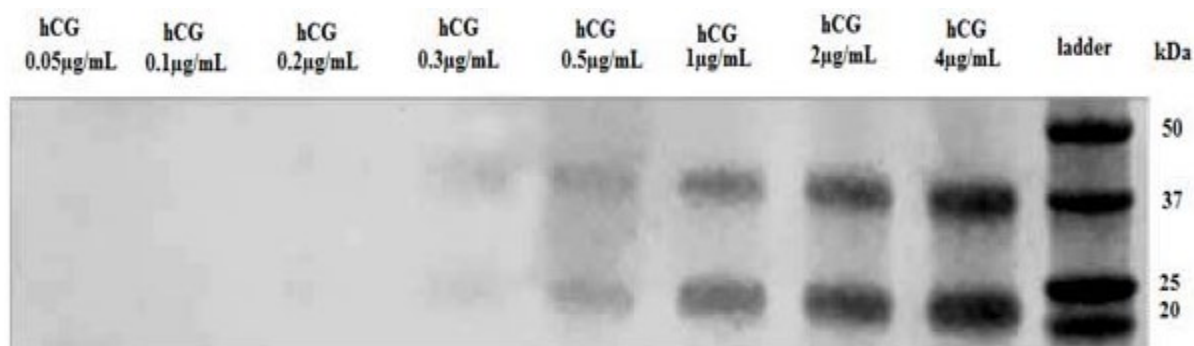

Figure 2.11 SDS-PAGE of eluted hCG standards (AC048) at concentrations of 0.05 to 4 µg/mL from Ab - Dynabeads using 0.1% TFA after 12 hours incubation on the rolling platform at room temperature.

2.4 Further optimisation of hCG enrichment with solid-phase extraction using Dynabeads®

The Dynabead enrichment method permitted the detection of hCG using SDS-PAGE down to 0.3 µg/mL after overnight incubation at RT. However, further optimisation of attachment conditions for antibodies to the magnetic Dynabeads and antibody-coupled beads to the hCG was optimized. Furthermore, hCG elution recovery from the beads as well as optimal sample matrix for further examinations using MALDI-TOF and hCG ELISA were studied.

2.4.1 Attachment of antibodies to magnetic Dynabeads

2.4.1.1 Methods

Table 2.5 shows the optimisation step for conjugated anti-hCG beta antibodies binding with magnetic tosylactivated Dynabeads over a longer time up to 64 hours. 100µg samples of anti-hCG beta (5004) and anti-hCG beta (5008) antibodies from Medix Biochemica (MedixMAB™), Finland were prepared and attached to the beads in a first instance (protocol Section 2.2.2.3). The efficiency of antibody attachment to the beads was measured over 12, 24, 48 and 64 hours on a roller at room temperature. At each time point, the protein concentration in the liquid phase was measured by placing the sample on the magnet for 2 min. 4µL was removed and protein concentration was determined using NanoDrop™ Lite Spectrophotometer (Thermofisher, UK) as presented in Table 2.6. This investigation established the mean percentage of antibodies bound to the beads over time.

2.4.1.2 Results

The effectiveness of antibody bounding to the beads was established based on the calculated amount of antibody remaining in the solution after precise incubation time (unbound fraction) (Table 2.5). After 12 hours of anti-hCG beta antibodies incubation with Dynabeads, the mean concentration of proteins in the supernatant was measured using NanoDrop. The mean percentage of total protein remaining in the solution after 12h was (18.82 ± 0.04mg/mL) which corresponds to the percentage of antibodies bound to the beads (81.17 ± 0.46). However, after 24 hours of incubation 86% of antibodies were bound to the beads. Longer incubation time did not significantly increase the amount of antibody bound which only slightly increased when the incubation time was 64h (+1.22%). Therefore optimal antibody to beads attachment was established at 24 hours and used in further experiments.

Table 2.5 Effect of incubation time on the binding of hCG antibody (0.1 mg in 1 mL diluent), from Medix Biochemica (MedixMAB™), Finland to magnetic tosylactivated Dynabeads at room temperature.

Incubation time (h)	Antibody (mg) in solution measured by NanoDrop (mean ± SEM)	Percentage (%) of antibody remaining in solution (unbound) (mean ± SEM)	Percentage (%) of the attached antibody to the beads (mean ± SEM)
0	0.102 ± 0.002	100 ± 0.25	<0.1 ± 2.53
12	0.019 ± 0.0004	18.82 ± 0.04	81.17 ± 0.46
24	0.014 ± 0.0004	13.69 ± 0.04	86.30 ± 0.39
48	0.014 ± 0.0006	13.93 ± 0.06	86.06 ± 0.61
64	0.013 ± 0.0008	12.71 ± 0.08	87.28 ± 0.79

2.4.2 Attachment and recovery of hCG to magnetic Dynabeads

2.4.2.1 Methods

Next, 50µL of the antibody-coupled beads were prepared as previously described (Section 2.2.2.3) and then incubated with hCG standard (1, 5ug; AC048, Fitzgerald, USA) in 1mL of 50mM AMBIC, pH 8, RT, on a roller for 0, 2, 4, 8 hours at room temperature. Following incubations, four washing steps with 500µL of PBS with 0.05% Tween 20, 500µL of PBS x 2 and 500µL of 50mM AMBIC, vortex mixing and removing the supernatant between each wash were undertaken. The hCG was eluted by the addition of TFA (15µL, 0.1%TFA, pH 2.5) following continuous shake (1000rpm) at room temperature for 10 minutes. Then, the solution was neutralised using 13mM ammonia solution (33.6µL, pH 10.3) and 50mM ammonium bicarbonate (20µL, pH 8.4) to obtain a mixture of beads and detached protein. Samples were then placed on the magnet (approximately 2 min) and the supernatant was transferred to a new collection tube. At each time point, the protein concentration in the liquid phase was measured by placing the sample on the magnet for 2 min. 4µL was removed and protein concentration was determined using NanoDrop™ Lite Spectrophotometer (Thermofisher, UK) as shown in Table 2.6.

2.4.2.2 Results

Incubation of antibody-coupled beads with 1μg of hCG protein standard (AC048, Fitzgerald) showed 100% attachment to the beads after 2 hours incubation at RT. When 5μg of hCG protein standard was added to the Dyneabeads, the percentage of hCG bound to the beads increased from 55% after 2 hours to 70 % after 4 hours incubation at room temperature. Extended incubation time up to 8h, resulted in a 5% drop in attachment Results in Table 2.6 show that 50μL of antibody-coupled beads bound 2.75 ± 0.48 μg (55 ± 9.57%) of hCG after 2 hours and up to 3.5 ± 0.28 μg (70 ± 5.77%) after 4 hours of incubation.

Table 2.6 hCG and antibody-conjugated Dynabeads attachment and recovery measured by NanoDrop™ Lite Spectrophotometer (ThermoFisher, UK). Time 0h is the 1 and 5μl hCG in solution immediately after adding the beads.

Incubation time (h)	Initial amount of hCG (AC048) standard (μg)	hCG (mg) in solution measured by NanoDrop after enrichment (mean ± SEM)	Percentage (%) of the attached hCG to the beads (mean±SEM)
0	1 μg	0.001	<0.1 ± 0.00
	5 μg	0.005	<0.1 ± 8.16
2	1 μg	0.000 ± 0.000	100 ± 0.00
	5 μg	0.002 ± 0.001	55.0 ± 9.57
4	1 μg	0.000 ± 0.000	100 ± 0.00
	5 μg	0.002 ± 0.001	70.0 ± 5.77
8	1 μg	0.000 ± 0.000	100 ± 0.00
	5 μg	0.001 ± 0.001	65.0 ± 9.57

2.4.3 Elution of hCG fromDynabeads

The above results showed 70% recovery using 5μg hCG for 4 hours. Therefore, it was shown that by using the optimized method with 50 μL magnetic beads, 3.5μg of hCG can be purified. Following manufacturer recommendation 50μg of the antibody-conjugated beads may typically bind to 1-10μg of protein (ThermoFisher, Invitrogen UK). As a consequence, the Dynabeads enrichment technique was investigated in order to further boost hCG recovery by focusing on the conditions of the final elution step (hCG separation from Dynabeads). This included analysis of an effect of eluent concentration on peak intensity using MALDI-TOF and hCG recovery after elution from Dynabeads.

2.4.3.1 Optimization method

Briefly, 50μL of the antibody-coupled beads prepared as previously described in Section 2.2.3.3 was incubated with 5μg of hCG protein standard (AC048, Fitzgerald) (1mL, 50mM AMBIC, pH 8, RT) on a roller for 4 hours. Samples were collected at each stage of the process Thus, four washing steps with 500μL of PBS with 0.05% Tween 20, 500μL of PBS x 2 and 500μL of 50mM AMBIC, vortex mixing and removing the supernatant between each wash were undertaken. The hCG was eluted by the addition of TFA (15μL, 0.1%TFA, pH 2.5) following continuous shake (1000rpm) at room temperature for 10 minutes. Then, the solution was neutralised using 13mM ammonia solution (33.6μL, pH 10.3) and 50mM ammonium bicarbonate (20μL, pH 8.4) to obtain a mixture of beads and detached protein. Samples were then placed on the magnet (approximately 2 min) and the supernatant was transferred to a new collection tube. In this experiment, elution steps were repeated three times by removing 128.6μL of supernatant and adding fresh 15μL TFA (0.1%), 33.6μL ammonia solution (13mM) and 20μL ammonium bicarbonate (50mM) each time. Next, to measure hCG concentration from each elution (1st, 2nd and 3rd), 1μL of supernatant was collected and diluted with reagent diluent at 1:1000, 1:10000 and 1:100000 following hCG beta ELISA R&D, UK procedure mentioned in Section 2.2.3.2. Subsequently, 3 x127.6μL, pH 9.4 were briefly vortex and left to dry in the vacuum concentrator (1.5h, 30°C). Later each sample was reconstituted in ultrapure water and digested using the procedure for Section 2.2.2.1 including denaturation of protein (boil, 10min) addition of 10mM DTT (10μL, 56°C, 10min), 20mM IAA (10μL, RT, 10min) and trypsin 20:1 ratio (20μg/mL, 37°C, 19h) accordingly, following incubation at 37°C, overnight. Lastly, each sample (1μL) was mixed with 1μL HCCA+DHB matrix (10mg/mL) and spotted onto a 384well target plate. MALDI-TOF analysis was carried out using the conditions previously described in Section 2.2.2.3. the procedure was also repeated with three elutions combined, dried in a concentrator (3h, 30°C) and spotted on the MALDI-TOF target plate as described above.

2.4.3.2 Results of extended elution conditions

Figure 2.12 demonstrates that by performing the elution procedure three times and by combining all three elution fractions (derived from 5g hCG (AC048) bound to Dynabeads) total hCG recovery increased by 16.7%. Accordingly, first, second and third elution measured contained 3.48 ± 0.107μg; 0.714 ± 0.022 and 0.149 ± 0.014, which converts to 69.7%, 14.3% and 2.9% hCG recovery from the beads. Thus after combining three elution steps, the final hCG recovery was 86.7% (in comparison to 70% (Table 2.6)) and an incubation time of 4h was

found to be most effective. hCG was also detected in the wash solution at 0.28±0.147 which was equally to 5.6% as shown in Figure 2.12. Furthermore, each collected sample from consecutive stages was analysed using ELISA and MALDI-TOF, as on chromatogram with results (Figure 2.13). However, it was observed that the fluctuations of hCG concentration in the samples were not proportional to the peak intensity observed on the chromatogram.
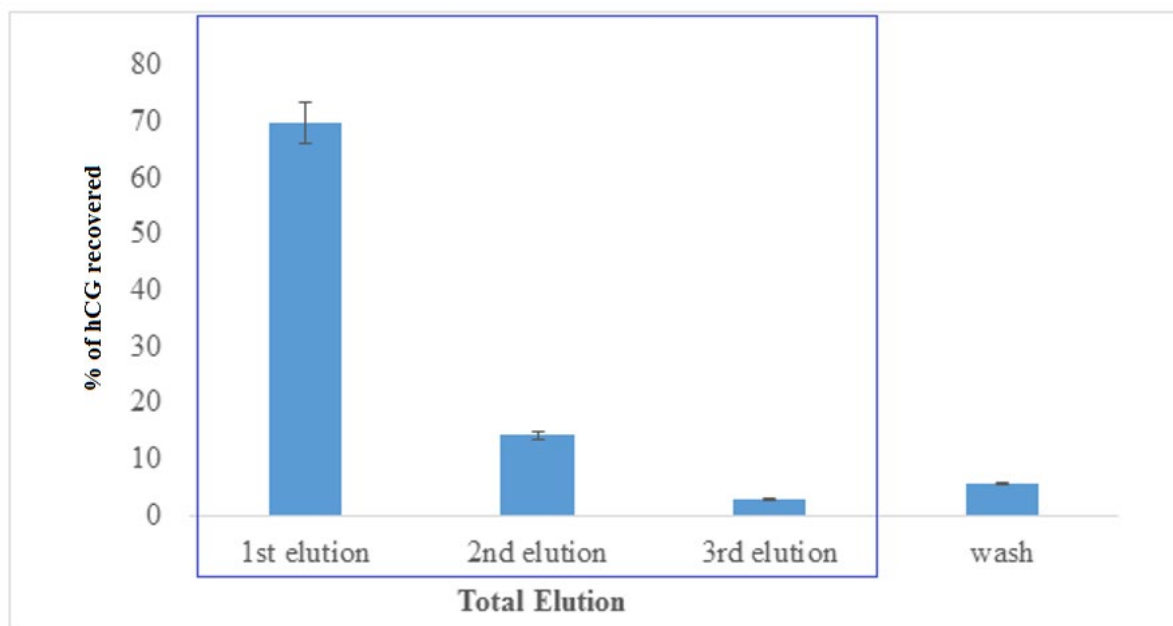

Figure 2.12 Percentage of recovered hCG in four consecutive stages during optimization of Dynabeads enrichment procedure (wash, first, second, and third elution) after 4 hours incubation at room temperature using hCG beta ELISA R&D, UK. Error bars represent the standard error of the mean (n=5) (SEM).
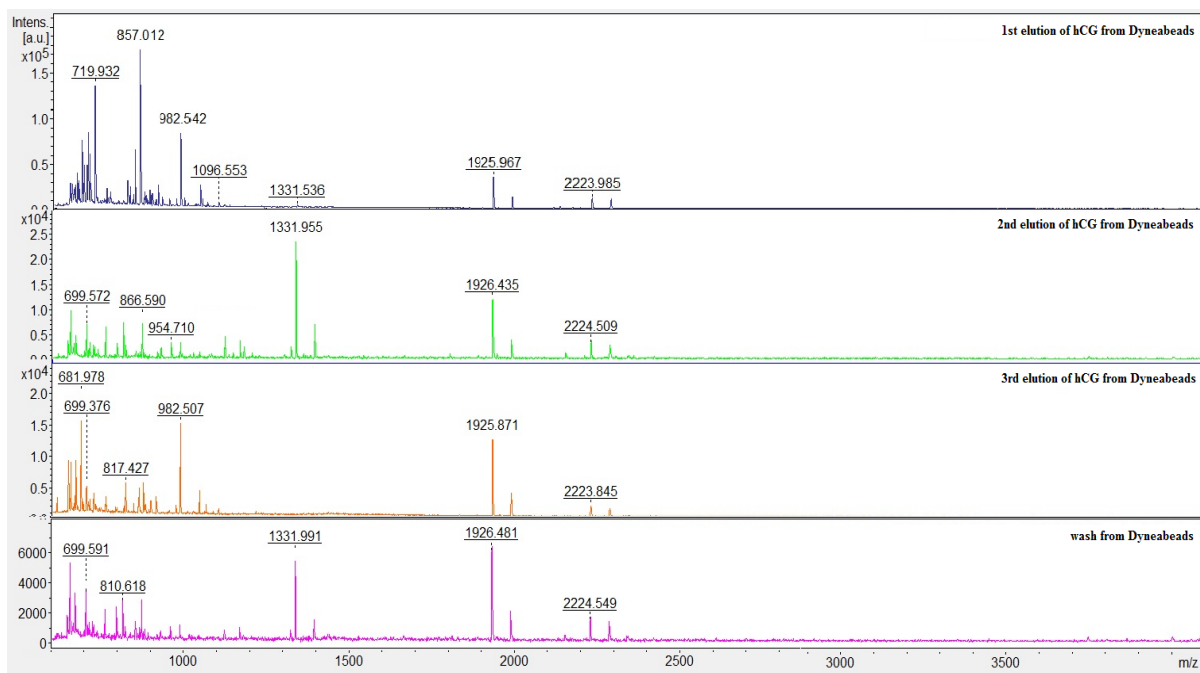
Figure 2.13 MALDI-TOF chromatogram of eluted and trypsinised hCG fractions following enrichment process using Dynabeads (4 h, RT). The four different fractions comprise first, second, third elution and wash, respectively.

Hence, the presence of hCG eluted from Dynabeads was confirmed in all four samples by comparing polypeptide sequence coverage from the samples to the sequence of human choriogonadotropin subunit beta variant, *CGB1, CGB7* and *CGB2* of the human genome using Mascot software on MALDI-TOF MS/MS. Furthermore, in order to fully identify specific picks in each eluted hCG sample from the Dynabeads as visible in Figure 2.13., Table 2.7 was constructed with information about m/z of prominent peaks obtained from control digestions of anti-hCG antibody, Dynabeads only, and standard hCG. Peaks present in eluted fractions were matched specific m/z peaks identified only in the hCG standard.

Table 2.7 Peaks (m/z) obtained after trypsinisation of the final eluent of hCG using Dynabeads (4h, RT). The alignment of m/z peaks found in Dynabeads elution corresponds to the known peptide sequence of hCGβ and IgG1 obtain using the Mascot engine and SwissProt database (expected relative errors ±500ppm). Non-relevant peaks from digestion i.e. trypsin, buffers, and matrix were included.

Peaks detected in eluent of hCG enriched by Dynabeads (m/z)	Products of digestion of Dynabeads (m/z)	Products of digestion of ant-hCG Antibody (m/z)	Products of digestion of hCG (AC048) (m/z)
643.450	643.792		
650.071	650.062		
651.396			651.396
656.086			656.062
666.030	666.012		
672.044		672.016	
681.990	681.979		
688.011	687.992		
692.079		692.059	
694.075	694.631		
697.954	697.943		
699.385			699.385
703.973	703.959		
709.968	710.695		
714.270		714.169	
719.945	719.424		
721.944			721.303
726.177	726.335		
742.249		742.300	
807.388	807.371		
817.436			817.435
841.032			841.031
842.489			842.258
857.012	857.012		
866.590	866.590		
876.987			876.773
892.950			892.744
908.916			908.706
954.710	954.710		
982.526			982.478
990.449	990.432		
1006.516	1006.491		
1096.553	1096.553		
1331.955	1331.955		
1925.921			1925.92
2224.892			2224.8919
2284.954			2284.9544
2468.032			2468.0325
2679.952			2679.9517
2905.142			2905.141
4116.975			4116.912

2.5 Application of Dynabeads to clinical samples

Subsequently, the developed Dynabeads enrichment method with 86.7% recovery of hCG after 4 hours incubation at RT was applied to enrich hCG from healthy pregnancy urine samples.

2.5.1 Enrichment of hCG from pregnancy urine

Briefly, 165µL of paramagnetic tosyl activated (solid-phase) Dynabeads were washed three times in 0.1M sodium phosphate buffer, pH 7.4 (1mL) and then resuspended in 0.1M sodium phosphate buffer containing 100 µg anti-hCG beta antibodies of either 95µL (1.052 mg/mL) 5004 or 19.6µL (5.1 mg/mL) 5008 in total of 150µL (see Section 2.2.2.3). Next, 100µL of ammonium sulphate (3M, dissolved in 0.1M sodium phosphate buffer) was added and antibody-conjugated beads (5004 and 5008 separately) were incubated 24 hours at room temperature on a rotating platform. Then, the supernatant was removed and beads were incubated with 1mL PBS with 0.05% Tween 20 at 37°C, 800 rpm for 2h. A final wash of beads with 2x1ml of PBST (PBS containing 0.01% Tween 20), vortex mixing between each wash, resulted in 250µL (5mg/mL) beads-antibodies complex used in the next step. 50µL of the antibody-coupled beads (1mg) were incubated with 1mL of a pregnant urine sample (nine samples from various participants (Section 2.2.1.3) from each trimester) on a roller for 4 and 24 hours at room temperature. Then, to elute hCG attached to Dynabeads, beads were washed twice with 500µL of PBS with 0.05% Tween 20, 500µL of PBS x 2 and 500µL of 50mM AMBIC, vortex mixing and removing the supernatant between each wash. Next, hCG was eluted by the addition of TFA (15µL, 0.1%TFA, pH 2.5) following continuous shake (1000rpm) at room temperature for 10 minutes. Next, the solution was neutralised using 13mM ammonia solution (33.6µL, pH 10.3) and 50mM ammonium bicarbonate (20µL, pH 8.4) to obtain a mixture of beads and detached protein (128.6µL, pH 9.4). Samples were then placed on the magnet (approximately 2 min) and the supernatant was transferred to a new collection tube. These elution steps were repeated three times by removing 128.6µL of supernatant and adding fresh 15µL TFA (0.1%), 33.6µL ammonia solution (13mM) and 20µL ammonium bicarbonate (50mM) each time and combining 1st, 2nd and 3rd elution in one tube. Thus, a total of 385.8µL, pH 9.4 was briefly vortex and left to dry in the vacuum concentrator (3h, 30°C). Lyophilised hCG samples were stored at -80°C for further qualitative and quantitative analysis including hCG ELISA Dynabeads recovery. Thus for ELISA study hGC samples were reqonstituted in 1mL of reagent diluent. This is the same starting volume of urine sample as used before the enrichment step.

2.5.2 Results of hCG enrichment from pregnancy urine

The average hCG (intact hCG, hCGβ subunit) concentration in human pregnancy urine collected from seven healthy female participants (n=27) was 1637.2 µg/ mL. Results in Table 2.8 show starting hCG concentration before the enrichment step and after enrichment with Dynabeads. Incubation of anti-hCG antibody-coupled Dynabeads with 1mL of human pregnancy urine from various weeks of gestation (4 h, RT) showed on average 27.6% recovery. However, extended incubation time up to 12h overnight, at room temperature, increased hCG recovery from the beads up to 76.7%.

Table 2.8 Enrichment of hCG glycoforms purified from healthy pregnancy urine samples using Dynabeads. hCG concentration before and after enrichment was measured using ELISA, R&D, UK.

Sample	hCG concentration before enrichment [µg/mL] (mean ± SEM)	hCG concentration with Dyneabeads **4hours** Incubation [µg/mL] (mean ± SEM)	hCG concentration with Dyneabeads **12hours** Incubation [µg/mL] (mean ± SEM)
1st trimester pregnancy (n=9)	1270.8 ± 433.5	349.3 ± 200.1	1003.0 ± 646.1
2nd trimester pregnancy (n=9)	1023.6 ± 581.8	285.3 ± 211.5	759.0 ± 563.0
3rd trimester pregnancy (n=9)	2617.1 ± 684.0	722.0 ± 220.8	2006.1 ± 616.0

2.6 Discussion

In this chapter, three methodologies were investigated for the purification/enrichment of human chorionic gonadotropin collected from human pregnancy urine and conditioned media of choriocarcinoma cancer cells. These employed methods based on binding of hCG according to their carbohydrate-binding properties (concanavalin A) or by using solid phase-conjugated antibodies (protein G sepharose and Dynabeads).

Initially, we examined a lectin affinity chromatography technique to purify highly diverse glycan-binding hCG glycoproteins of non-immune origin. This technique has been used before as an effective tool for the enrichment of hCG glycoforms from urine samples and screening for potential cancer biomarkers (Matsuura, and Chen, 1980; Blithe, 1990; Koistinen *et al*., 2015; Hashim, *et al*., 2017). Lectin-based isolation/enrichment technique based on capturing glycans commonly uses concanavalin A (ConA) which selectively recognizes a diversity of carbohydrates with high-affinity binding to the mannose core in N-glycans and binds to them through hydrogen bonds, van der Waals and hydrophobic forces (Abbott and Pierce, 2010; Raposo, *et al.,* 2021). Figure 2.8 presented samples of biological samples including cultured media (JEG-3) and urine samples incubated with Con A (104kDa) and separated on the SDS-PAGE gel however, we did not observe any bands corresponding to eluted hCG in any of the lanes, which indicated a lack of carbohydrate attachment to the lectin. This might be caused by the low overall concentration of hCG and high dilution in the tested urine and cultured media samples (hCG ≤0.5mg).

Therefore, further investigation of more sensitive (for detection <1µg of hCG), and precise methods based on solid-phase extraction using conjugated antibodies beads with site-specific anti-hCG antibodies (protein G sepharose beads and paramagnetic Dynabeads) were examined. The anti-hCG monoclonal antibodies used to capture hCG, were specific to both the β subunit and αβ heterodimer simultaneously, as per advice from Prof Ulf-Hakan Stenman from University Central Hospital in Helsinki, Finland, and Lund, et al (2014). Moreover, used antibodies from Medix Biochemica (MedixMAB™), Finland, 5008 SP-5, and 5004 SP-1 can recognise a wide variety of hCG and hCG variations (pan hCG-mAbs), including intact hCG, nicked hCG (hCGn), hCGβ, hCGβn and hCGcf with 1% hLH cross-reactivity. Used in our experiments variant-selective antibodies (β2), were approved during the Second International Workshop on hCG of the International Society of Oncology and Biomarkers Tissue Differentiation 7 (ISOBM TD-7) as recognizing β1–β5 epitopes among two antigenic domains: the cystine knot (epitope β1) and hCGβ loops 1+3 comprising the neighbouring

epitopes β2–β6 (Berger, *et al*., 2013). Therefore, the choice of the above antibodies and conditions for solid phase-conjugated antibody beads aimed enrichment of a wide range of hCG glycoforms.

As a result of enrichment with agarose-sepharose protein G beads with a high affinity for mouse IgG antibodies, we were able to capture hCG; nevertheless, the amount of hCG recovered was quite low. The concentration of antibody used was within the recommended range i.e. 2mg/mL antibody to 1mL sepharose beads (Kavran and Leahy, 2014). Achieved results suggested poor binding of the hCG antibody to the protein G beads (Figure 2.9). Therefore, the coupling efficiency of the ligand with protein sepharose G (cyanogen bromide (CNBr)-activated) was difficult to accurately calculate. Additionally, a significant amount of unbound hCGs was found in wash fractions (Figure 2.10 (a)). Accordingly, this method allowed recovery and detection of up to 0.5μg/mL of hCG using SDS-PAGE (Figure 2.10 (b)). Protein G enrichment is typically used for purifying a high scale (mg) of proteins on columns from around 50 litres of urine hence, it was not practical to scale up our experiment (Khademi *et al*., 2009). Therefore, we tested separation methods using Dynabeads with immobilised antibodies against hCG with a lower concentration of starting material (0.3μg/mL of standard hCG). M-450 Tosylactivated Dynabeads uniform, superparamagnetic polystyrene beads (4.5 μm diameter) with reactive sulphonyl esters were used, which can covalently bind with mixed antibodies By optimising pH, temperature and time of incubation, beads were binding to antibodies (or standard proteins) effectively, hence the same conditions were used with urine samples. As a result of the optimized Dynabeads method, hCG was extracted from pregnant urine with 76.7% efficiency after 12 hours of incubation. This immunocapture approach allowed effective enrichment from complex samples for low abundant proteins and allowed hCG purification from biological samples (Rodriguez*, et al*., 2005; Lund *et al*., 2013; Levernæs *et al,* 2017). Other popular methods for hCG isolation from urine samples include affinity chromatography, ion-exchange and size-exclusion chromatography, gel electrophoresis or enzyme-linked immunosorbent assay (Xu, *et al*., 1999, Melchor et al., 2005, Khademi, *et al*., 2009, Rosen *et al*., 2013). Typically, these methods require high volumes of urine ≥50mL and are based on the specificity and cross-reactivity of anti-hCG antibodies, which are unable to detect and quantify a variety of hCG isoforms drastically fluctuating throughout pregnancy and pregnancy disorders. Therefore, it is very challenging to accurately, instantaneously and equally detect all hCG isoforms using currently available biochemical tests (Berger, *et al.,* 2013). Despite the fact, that the tests are very sensitive (~5 IU/L), many different hCG isoforms such as free β subunit, nicked free β subunit, c-terminal peptide, β-core, or hyperglycosylated hCG have to

be tested independently (Gnoth and Johnson, 2014, Berger and Lapthorn, 2016). As a result, concerns regarding the use of commercial diagnostic immunoassays for the diagnosis of GTDs or pregnancy disorders include uneven identification and detection of the various hCG glycoforms, particularly hCGβcf due to a lack of standardised diagnostic antibodies and in particular antibody epitopes' that might alter the final test findings. Additionally, heterophilic antibodies that are generated by exogenous antigens against antigens that are not well defined might bind to already captured and labelled antibodies resulting in false-positive outcomes. However, they are only detected in serum samples and are not detectable in urine hCG testing (filtered in the kidneys) (Esfandiari and Goldberg, 2003, ACOG, 2002, Bolstad and Warren, 2013). Accordingly, this all leads to increased likelihoods of errors, irreproducible outcomes and false negative or false positive readings (Mangili *et al.,* 2014, Seckl *et al*., 2010, Brown *et al*., 2017). Thus, comparative epitope mapping of diagnostic epitope-defined reference antibodies will increase the reproducibility of hCG measurements; nevertheless, these procedures are still underway (Berger and Lapthorn, 2016). Therefore, there is a clear indication that by optimising and enhancing proteomics and glycomics detection methods, and minimising the reliance on antibody usage, we could advance very powerful diagnostic tests to detect hCG glycoforms identified as biomarkers for pregnancy disorders or cancer.

In summary, we have developed an efficient and precise method for hCG purification using Dynabeads with 76.7% recovery when applied to clinical samples. This method allowed to further investigate the glycan heterogeneity of hCG in clinical samples including normal pregnancy urine and urine collected from gestational trophoblastic diseases including choriocarcinoma.

3. Development and optimisation of the method for hCG glycans characterization

In the preceding chapter, the development of a method for enriching and purifying hCG glycoprotein using antibody-conjugated Dynabeads was described. To detect and characterise N-linked glycans present on hCG and other glycoproteins using MALDI-TOF MS, carbohydrates were removed from the parent protein through a process known as deglycosylation. The initial experimental design was based on Valmu *et al.* (2006), Jang *et al.* (2009), and Levernæs *et al.* (2017) methods, as well as personal discussions with Dr Thomas Abban (Shimadzu) and Dr Julia Smith (Bruker) who are holding expertise in proteomics and glycomics studies. Ovalbumin as a model glycoprotein with well-characterised covalently N-linked glycans was chosen in preliminary studies to optimise conditions for deglycosylation and detection of N-linked glycans. Results were compared to glycans moieties described in the Glycobiology Analysis Manual, (Sigma Aldrich, 2016) and Harvey *et al.* (2000). This chapter describes how each stage of the deglycosylation process, including heat denaturation, disulphide bond reduction, alkylation, and enzymatic glycan removal using PNGase F, was optimised (Section 3.2). The overall workflow is presented below in Figure 3.1. Released from a protein glycans were extracted using solid-phase extraction tips including HILIC, C18, and Porous Graphitic Carbon (PGC) in order to detect the highest number of N-glycans from the parent protein (Section 3.3) by MALDI-TOF mass spectrometry. Additionally, this chapter presents how sample preparation was optimized, instrument specifications and limitations, and optimal experimental settings throughout the experiment since all of those aspects had a significant impact on the sensitivity and accuracy of the glycans measurements after detachment from a protein. Lastly, glycans spectra were acquired using MALDI-Time of flight mass spectrometry in reflectron, positive ion mode over a range of m/z=900 Da – 4000 Da. N-linked glycans were identified by GlycoQuest software as [M+Na+, M+K+, M+Li+] ions with multiple modifications in glycan assembly (mono-, bi-, tri- and tetra- antennary or high mannose) and/or composition such as fucosylation, sialylation and a substantial amount of truncated, bisected and sulphated glycans detected. In further steps achieved results from MALDI-TOF (Bruker, UK) were also compared with LC-MS/MS QExactive Orbitrap (ThermoFisher, UK) which enabled the verification of matched glycans based on the closeness of measurements agreed by two methods.
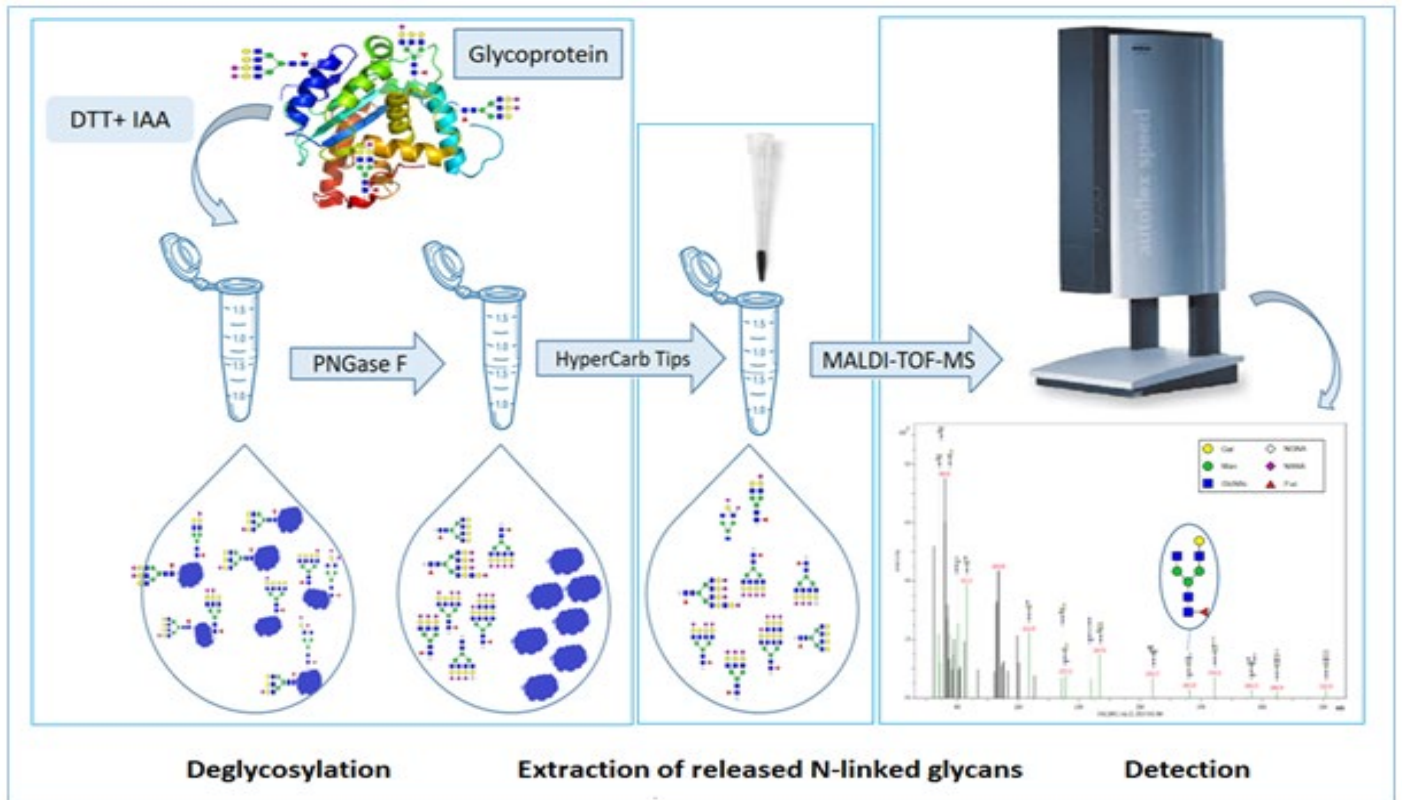
Figure 3.1 Schematic diagram of the method development process for N-linked glycans characterisation (Adapted from Gondek, *et al*., 2018).

3.1 Materials and Reagents

3.1.1 Major and minor equipment

Centrifuge	Sigma Aldrich, UK
Heating Block	ThermoFisher, UK
Balance	ThermoFisher, UK
Vacuum concentrator	Eppendorf, UK
Vortex mixer	ThermoFisher, UK
Pipettes	Brand, Germany
Thermomixer	Eppendorf, UK
MALDI IT-TOF MS/MS (Autoflex)	Bruker, UK
Flex Analysis – GlycoQuest software	Bruker, UK
GycoWorkBench software	Informer Technologies, Inc.UK
LC-MS/MS QExactive Orbitrap	ThermoFisher, UK

3.1.2 Reagents and chemicals

- Commercially available hCG standards all purified from human urine: intact hCG (AC048), Fitzgerald Laboratories, USA and intact hCG (C0434, Batch 1 and 2), Sigma, UK
- hCG collection from human pregnancy urine of seven healthy participants (4-33 weeks) collected between November 2016 to March 2019
- 2,5-Dihydroxybenzoic acid (2,5 DHB- 85707) matrix substance for MALDI-MS, >99.0% (HPLC), Sigma Aldrich, UK
- Alpha-Cyano-4-hydroxycinnamic acid (4-HCCA-70990) matrix substance for MALDI-MS, 99.0% (HPLC), Sigma Aldrich, UK
- Dithiothreitol (DTT - D0632) ≥98% (HPLC), ≥99.0% (titration), Sigma Aldrich, UK
- Iodoacetamide (IAA - A3221) ≥99% (HPLC), Sigma Aldrich, UK

- Trypsin from porcine pancreas, Sigma Aldrich, USA, Cat No T6567, Lot 031M6160V, for MALDI analysis was prepared by reconstituting the lyophilized product in 50 mM Ammonium bicarbonate and 10% acetonitrile (stock solution 100ng/µl)

- Precision Plus Protein Standards Dual Colour Ladder, 161-0394, Bio-Rad, UK

- HILIC HyperSep™ Tip Microscale SPE Extraction Tips, 1-10µL (60109-206), ThermoFisher, UK

- BioBasic C18 HyperSep™ Tip Microscale SPE Extraction Tips, 1-10µL (60109-201), ThermoFisher, UK

- HyperSep Hypercarb Porous Graphitic Carbon (PGC) Microscale SPE Extraction Tips, 1-10µL (60109-204), ThermoFisher, UK

- 4–15% Mini-PROTEAN® TGX (4561083SEDU), BioRad, UK, precast polyacrylamide gels with a 10-well comb

- Sample buffer (2x Laemmli, dilute 1:1 with sample) 62.5 mM Tris-HCl, pH 6.8, 2% SDS, 25% glycerol, 0.01% bromophenol blue

- Running buffer (Tris/glycine/SDS) 25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3

- Phosphate Buffered Saline (PBS) tablets, BR0014, Oxoid, UK; formula: 8g sodium chloride, 0.2g potassium chloride, 1.15g disodium hydrogen phosphate, 0.2g potassium dihydrogen phosphate were dissolved in nuclease-free water, pH7.3

- Trifluoroacetic acid (TFA) 99+%, for HPLC (10294110), Fisher Chemical, ThermoFisher, UK

- Nuclease-free water (10429224), UHQ, Invitrogen™, ThermoFisher, UK

- 50 mM Ammonium bicarbonate (AMBIC) was prepared by adding 79mg of AMBIC (10207183) which is 99% pure powder used for analytical analysis purchased from ACROS Organics™, (ThermoFisher), UK to 20mL of Nuclease-Free Water (10429224), UHQ, (Invitrogen™ThermoFisher, UK)

- 0.1% trifluoroacetic acid (TFA- T62200), Sigma Aldrich, UK, was prepared by adding 200µL of TFA in 200mL of distilled water

- Acetonitrile (ACN) for HPLC-MS (10616653), ThermoFisher, UK

- Glycan NGA2F (CN-NGA2F-x) purified from the oligosaccharide pool released from porcine thyroglobulin by hydrazinolysis using a combination of HPLC and glycosidase digestion (Ludger, UK)

- Peptide-N-Glycosidase F also known as PNGase F (peptide-N-(acetyl-β-glucosaminyl) asparagine amidase. This was cloned from *Flavobacterium meningosepticum* and expressed in *E.coli 3.2.218; 3.5.1.52*); PNGase F (P0704S) from New England Bio Labs (NBS), UK the one unit is defined as the amount of enzyme required to remove > 95% of the carbohydrate from 10 µg of denatured RNase B in 1 hour at 37°C in a total reaction volume of 10 µL (1 activity unit/µL) and PNGase F (11365177001), Roche, Welwyn Garden City, UK. 3-300 units/mL deglycosylate up to 500 µg of glycoprotein/mL (1.66 activity unit/µl).

3.2 Optimalisation of a protein deglycosylation

Deglycosylation of a selected protein was achieved by the actions of temperature, dithiothreitol (DTT) iodoacetamide (IAA) and the activity of peptide-N-glycosidase (PNGase F). The initial denaturation was induced by heating of glycoprotein (100°C for 10 min) which allowed to disturb the secondary and tertiary protein structure, hydrogen bonds and non-polar hydrophobic interactions. High temperature increases kinetic energy and excites the molecules to rapidly and violently vibrate, which unravelled chains in a protein into a random shape (Matsuura, *et al.*, 2015). Dithiothreitol (DTT) is a redox reagent with reducing and oxidising properties. DTT reduces the disulphide bonds of proteins by losing an electron, hence its oxidation state increases. DTT prevents bonds from forming between cysteine residues of proteins, especially under denaturing conditions (O'Neil, 2001). The alkylating agent; 2-Iodoacetamide (IAA) was used to bind covalently with the thiol group of cysteine, so the protein cannot form the disulphide bonds again (Buskas *et al.,* 2006). Next, the enzymatic removal of glycans was carried out using Peptide-N-Glycosidase F (PNGase F). Deglycosylated proteins were analysed using MALDI-TOF MS; GlycoWorkBench and GlycoQuest software. In the following sections, the effect of temperature, DTT and IAA concentration are described before a step of PNGase F treatment which effectiveness was evaluated when ovalbumin and hCG (C0434) samples were used.

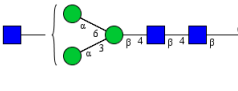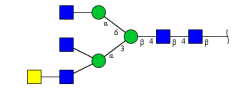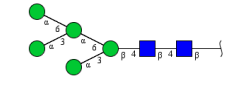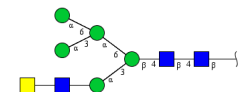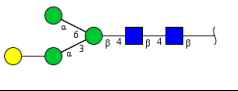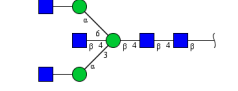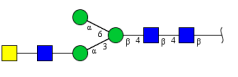### 3.2.1 Effect of temperature, DTT and IAA on deglycosylation
Method
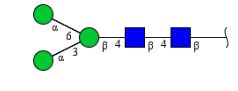
Ovalbumin (10µg in 10µL, 50mM AMBIC, pH 8) was incubated at room temperature or 100°C for 10 minutes, before the addition of 10 µL DTT (1 - 50 mM in ammonium bicarbonate (AMBIC) or 10µL AMBIC only. Samples then were incubated at 56°C for 15 minutes (Shevchenko, *et al*., 2006) prior to the addition of 10µL IAA (20 or 100 mM iodoacetamide in AMBIC) or 10µL AMBIC only. Mixtures were incubated at room temperature for 15 minutes and then 10µL of PNGase F (P0704S, New England Bio Labs,) was added to each sample. Samples were gently vortexed and incubated at 37°C overnight. Next, the resulting glycan mixtures were loaded onto solid-phase extraction tips HyperCarb Porous Graphite Carbon (PGC) and aspired 10 times in order to retain all the non-polar molecules (deglycosylated proteins) and elute glycans, after the addition of 0.1% TFA in 50% ACN. The eluent was dried in a vacuum centrifuge (30°C for 6h) reconstituted in 2µL of UHQ water, mixed (1:1) with matrix mix of 10mg/mL 2,5-Dihydroxybenzoic acid (DHB) and 10mg/mL Alpha-Cyano-4-hydroxycinnamic acid (HCCA) in 50% ACN/0.1% TFA and spotted onto a 384 well steel

MALDI target plate. MALDI-TOF MS parameters were as follows: reflector positive mode with m/z range of 700-5000Da and 6_ultra smartbeam parameter set with the zoom range of ±5%. Additionally, Mascot search parameters included SwissProt database, NCBI BLAST taxonomy browser limited to *Homo sapiens* species. Protein view: CGB7_Human, Chorigonadotropin subunit beta 7 OS=Homo sapiens OX=9606 GN=CGB7 PE=2SV=1. Experimental mass values were set to monoisotopic and variable modifications such as carbamidomethyl (C) and oxidation (M). The error window on the experimental peptide mass value was set to 500ppm. Mass values corresponded to charge carrier $[M+H]^+$. The number of missed trypsin digest cleavages was set to 1, which is optimal for founding the target protein against the experimental data. The GlyoQuest database search method, version 1.3, was used to analyze glycan data as part of the Bruker ProteinScape database programme (version 3.0). The GlycoQuest matched m/z MS or MS/MS data against theoretical spectra calculated from databases entries. It worked by importing MS data (glycan pool spectra) and selecting a range of search parameters such as type of glycan (hierarchical classification as defined in GlycomeDB database). The repertoire of N-linked glycans was obtained from public databases synchronized with ProteinDB, as well as a customised internal database (Search method: MG Glycan MALDI MS AA Carb, Method name: MG Glycan MALDI MS AA CarbBank) developed in Proteinscape and utilised for GlycoQuest (glycan) searches. Glycan composition (only glycan structures that included the chosen kinds of monosaccharides were considered in the search), i.e. only N-Glycan core basic (minimum 3dXex+Hex2NAc). Modifications (e.g. sulphate, phosphate) include Sulphate (0-4), N-Sulphate (0-1), derivatization type (various labels) yet, in this study all glycans were underivatized. The ions (neutral) attached (ions defined as adducts providing a charge to the ions detected) i.e. H+(0-1), Na+(0-2), K+(0-1), Li+(0-1). The above parameter settings were used in a search that included neutral exchanges H+ vs. Na+, which were common in acidic glycans (exchange of acidic proton in sialic acid by Na+). Finally, the MS tolerance (m/z) detection was set to (m/z) 0.7 Da and monoisotopic peaks were detected.

Results

The effect of temperature and addition of 20mM DTT and 100 mM IAA on ovalbumin is shown in Table 3.1. Incubation at room temperature with DTT and IAA, and no denaturation step yielded the removal of two N-linked glycans. Subsequently, 100ºC incubation in the absence of DTT and IAA allowed the detection of three glycans. When the sample was heated to 100°C degrees with DTT only, four glycans were visible. In contrast, when samples were heated with IAA only, 5 glycans were detected. Nevertheless, the most effective conditions involved heating samples to 100 ºC followed by incubation with both DTT and IAA which allowed the detection of eight N-linked glycan species (Condition 5 in Table 3.1). All detected N-linked glycans shared a common core sugar sequence, Manα1-6 (Manα1-3) Manβ1-4GlcNAcβ1-4GlcNAcβ1-Asn-X, in which only mannose (hexose) residues are attached to the core, by N-acetyl-glucosamine (GlcNAc). In Table 3.1, glycan composition is shown e.g. Hex3HexNAc2 – representing three Man and two GlcNAc.

Table 3.1 Effect of temperature (100°C or RT), DTT, IAA and PNGase F (in all samples) on the number and structure of detected glycans from ovalbumin. Monoisotopic glycan masses were measured in the positive ion reflectron mode as [M+Na$^+$] adducts, RT – room temperature

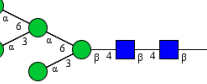| Glycan Structure	Glycan m/z & composition	Condition 1 RT DTT IAA	Condition 2 100°C - -	Condition 3 100°C DTT -	Condition 4 100°C - IAA	Condition 5 100°C DTT IAA
[glycan structure]	933 Hex3HexNAc2	+	+	+	+	+
[glycan structure]	1136 Hex3HexNAc3	+		+	+	+
[glycan structure]	1745 Hex3HexNAc6		+		+	+
[glycan structure]	1257 Hex5HexNAc2		+	+		+
[glycan structure]	1663 Hex5HexNAc4			+		+
[glycan structure]	1095 Hex4HexNAc2				+	+
[glycan structure]	1542 Hex3HexNAc5				+	+
[glycan structure]	1339 Hex3HexNAc4					+

3.2.2 Effect of DTT concentration on the glycan composition
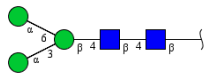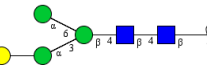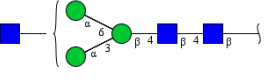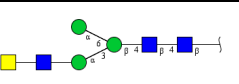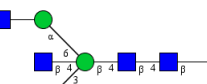
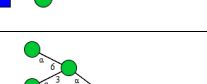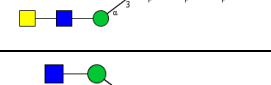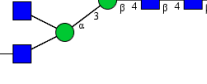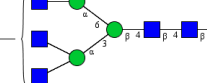**Method**

Ovalbumin (10µl, 1µg/µL in 50mM AMBIC, pH 8), was incubated at 100 °C for 10 minutes before the addition of various concentrations of DTT (1 - 50 mM, 10µL in AMBIC). Samples were then incubated at 56°C for 15 minutes. 10µL of IAA (100 mM in AMBIC) was then added to each sample and incubated at room temperature for 15 minutes. Next, 10µL of PNGase F (P0704S) from New England Bio Labs (NBS), UK was added to each sample. Samples were gently vortexed and incubated at 37°C overnight. Conditions are summarised in Table 3.2. Next solid-phase extraction tips HyperCarb Porous Graphite Carbon (PGC) were used. As mentioned previously, the glycan mixture was loaded onto the tips eluted (0.1% TFA in 50% ACN), dried (30°C for 6h), reconstituted in 2µL of UHQ water, mixed with 10mg/mL HCCA/DHB matrix and spotted (1:1) on the MALDI target plate. MALDI-TOF MS was set in a positive, reflector mode (m/z range of 700-5000Da) and glycans were characterised using the GlycoQuest software.

Results

The effect of various DTT concentrations on the ovalbumin is presented in Table 3.2. Glycans were detected from samples incubated at each DTT concentration. However, incubation of ovalbumin with 10 and 20 mM DTT each resulted in the detection of eight glycans. The number of glycans detected was decreased when other concentrations of DTT were used. Glycans that were detected ranged in size (m/z) from 933.20 Da - 2151.74Da. Structural analysis from glycan software revealed, as mentioned previously, that the smallest structure (m/z 933) comprised a core glycan structure made up of two N-acetyl glucosamine (GlcNAc) and three mannose residues [Hex3HexNAc2]. This is the precursor of all N-linked carbohydrate structures as presented in Table 3.2. Additionally, Figure 3.2 shows the chromatogram produced from a sample exposed to 10 mM DTT and 100 mM IAA showing eight peaks corresponding to the glycans listed in Table 3.2

Table 3.2 Effect of various DTT concentrations (0mM – 50mM) on amount of glycans detected from ovalbumin with specific m/z detected. Samples were denatured, 100 mM IAA was added and the composition of detected glycans corresponds to N-linked glycans from ovalbumin treated with PNGase F and spotted using HCCA matrix in a positive ion MALDI-TOF mode with M+Na$^+$ adducts.

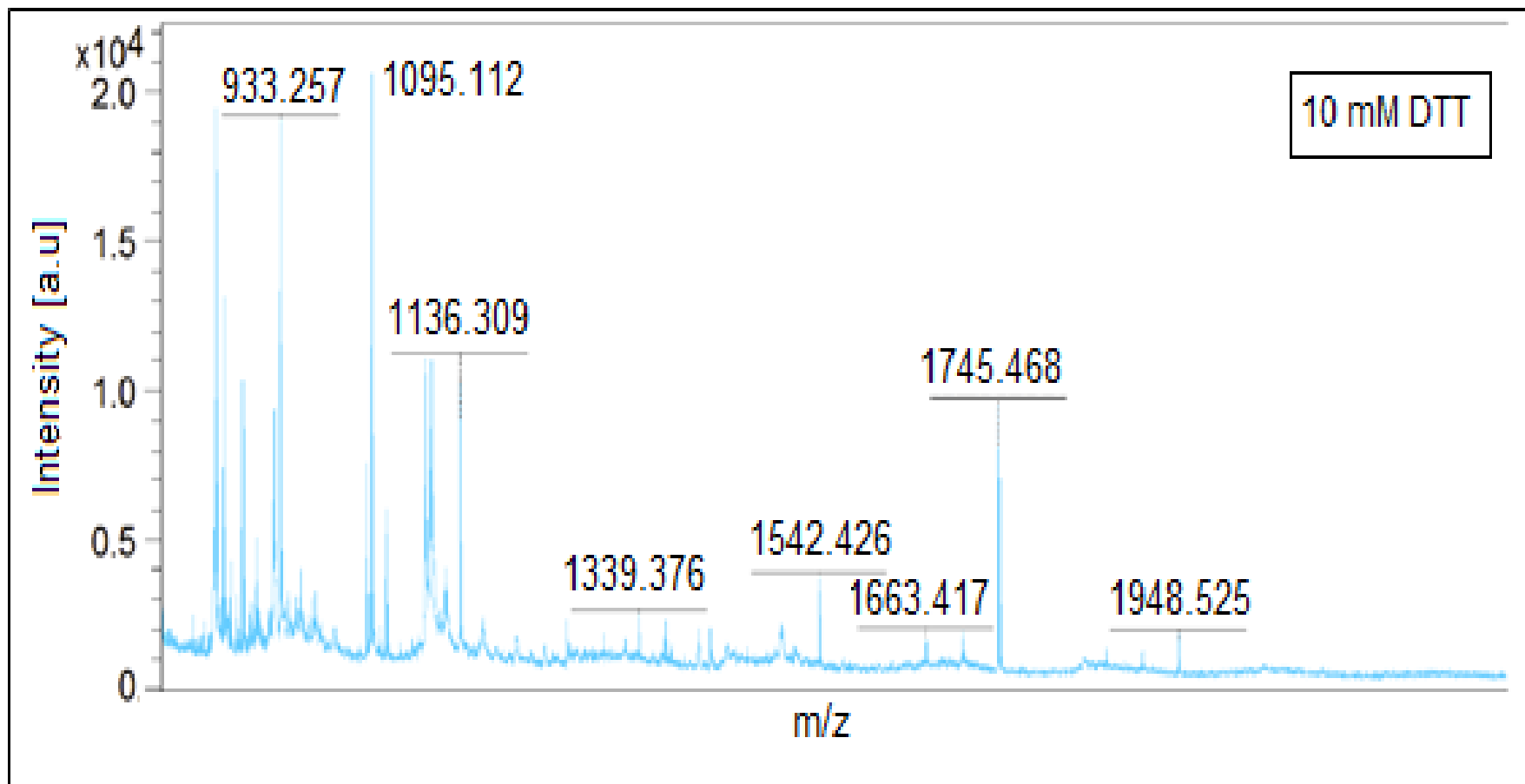| Glycan Structure	Glycan m/z & composition	DTT Concentration					
		0 mM	1 mM	5 mM	10 mM	20 mM	50 mM
	933 Hex3HexNAc2	+	+	+	+	+	
	1095 Hex4HexNAc2	+	+	+	+	+	+
	1136 Hex3HexNAc3				+	+	
	1257 Hex5HexNAc2		+			+	
	1339 Hex3HexNAc4				+	+	
	1542 Hex3HexNAc5				+	+	
	1663 Hex5HexNAc4				+	+	
	1745 Hex3HexNAc6			+	+	+	
	1866 Hex5HexNAc5						
	1948 Hex3HexNAc7				+		

Figure 3.2 Positive ion MALDI- TOF MS of the N-linked glycans from 10µg of ovalbumin released by 10µL of 10mM DTT, and 10µL of 100 mM IAA and treated with PNGase F. Following solid-phase extraction tips (HyperSep Carbon18) and mix with HCCA matrix. The mass spectra represents eight glycans detected from ovalbumin. Glycans m/z and composition are indicated in Table 3.2.

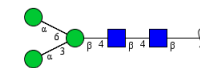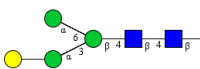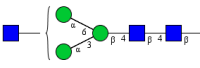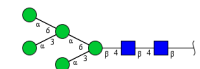### 3.2.3 Effect of IAA concentration on deglycosylation
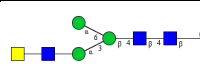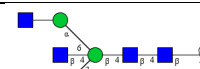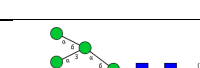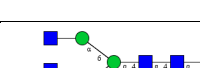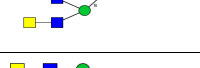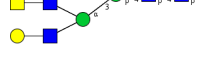
Method

Ovalbumin (10µg dissolved in 10µL of 50mM AMBIC, pH 8), was incubated at 100°C for 10 minutes before the addition DTT (10µL, 10mM in AMBIC). Samples were then incubated at 56°C for a further 15 minutes. Next, 10µL of IAA (20 or 100mM in AMBIC) was added and samples were incubated at room temperature for 15 minutes followed by incubation with 10 µL of PNGase F (P0704S) (New England Bio Labs (NBS), UK), gently vortexed and incubated at 37°C overnight. Next solid-phase extraction tips (HyperCarb Porous Graphite Carbon (PGC) were used to extract glycans from a mixture of deglycosylated glycoprotein. Thus, the glycan mixture was loaded onto the tips and only glycans were eluted (0.1% TFA in 50% ACN), dried (30°C for 6h), reconstituted in 2µL of UHQ water, mixed with 10mg/mL HCCA/DHB matrix and spotted (1:1) on the MALDI target plate. MALDI-TOF MS was set in a positive, reflector mode (m/z range of 700-5000Da) and characterised by using the GlycoQuest software.

Results

The effect of IAA concentration on the number of detected glycans is presented in Table 3.3. Both, the number of glycans detected and the intensity of their peaks were increased following incubation with 20mM IAA, compared to 100 mM IAA (Figure 3.3. (b)). In total twelve glycans (m/z 933.20 Da - 2151.74 Da) were detected in the sample with 20mM IAA. The chromatogram (Figure 3.3 (a)), shows the final effect of optimised experimental conditions for deglycosylation of 10µg ovalbumin, using heat (100°C), DTT (10mM in AMBIC) and IAA (20mM in AMBIC).

Table 3.3 Twelve individual N-linked glycan structures and compositions from ovalbumin (A-L) were detected with a sample where 20 and 100 mM IAA was added. The values of peak intensity correspond to the intensity of peaks detected on the chromatogram $[M+Na^+]$ (MALDI-TOF MS) (not corresponding to Figure 3.3 below).

Letter m/z peak	Glycan Structure	Glycan m/z & composition	20 mM IAA Presence of glycan	100 mM IAA Peak intensity
A		**933** Hex3HexNAc2	+	+
B		**1095** Hex4HexNAc2	+	+
C		**1136** Hex3HexNAc3	+	+
D		**1257** Hex5HexNAc2	+	+
E		**1339** Hex3HexNAc4	+	+
F		**1542** Hex3HexNAc5	+	+
G		**1663** Hex5HexNAc4	+	+
H		**1745** Hex3HexNAc6	+	+
I		**1866** Hex5HexNAc5	+	+
J		**1948** Hex3HexNAc7	+	+
K		**2110** Hex4HexNAc7	+	
L		**2151** Hex3HexNAc8	+	

(a)
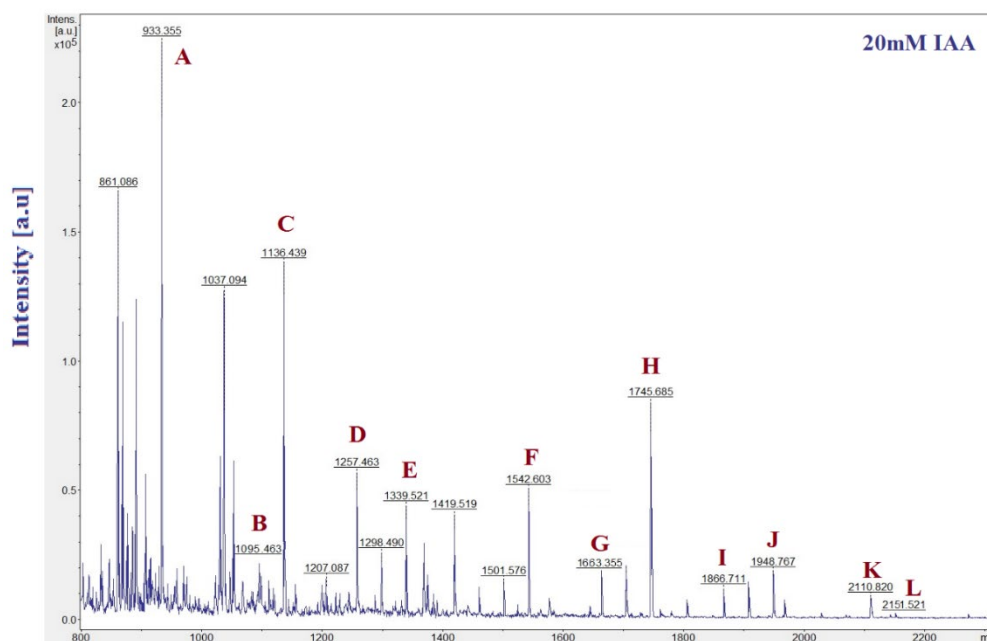
(b)
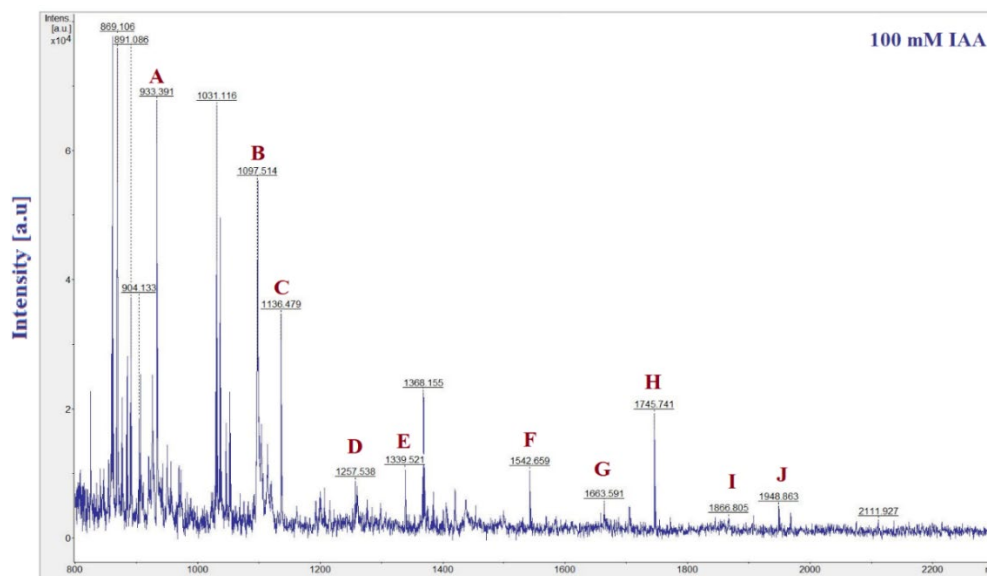
Figure 3.3 Positive ion MALDI- TOF MS of the N-linked glycans from 10μg of ovalbumin released by 10μL of 10mM DTT with 10μL of 20mM IAA (a) and 100mM IAA (b), following treatment with PNGase F, solid-phase extraction tips (HyperSep Carbon18) and mixed with HCCA matrix. Twelve and ten glycans structures were observed in samples with 20mM and 100mM IAA respectively. Note the different y- scales (up to 10^5 with 20mM IAA and 10^4 with 100mM) and the corresponding reduced noise was used.

3.2.4 Effect of DTT, IAA and PNGase F on deglycosylation

To this point, our findings indicated that the most successful deglycosylation process involved protein denaturation at 100°C for 10 minutes, followed by the addition of 20mM DTT (56°C, 15 minutes) and 20mM IAA (RT, 15min). These conditions yield the highest number of N-linked glycans detectable by MALDI TOF analysis. In the next step, the project aimed to compare the effectiveness of PNGase F which cleaves precisely between the GlcNAc and asparagine residues of N-linked glycoproteins and glycopeptides. Two commercially available PNGase F purchased from New England Biosciences (NBS), UK and Roche, UK were compared. The NBS, UK PNGase F unit definition required 10 μL of PNGaseF (65NEB units=1IUB miliunit) for removal of ≥95% of the sugars from 10μg protein and Roche, UK assay conditions required up to 300unit/mL of PNGase F for deglycosylation of up to 500μg of glycoprotein. In addition to both used PNGases F, the effect of DTT (reduction step) was examined (Figure 3.5). MALDI-TOF and SDS-PAGE were used to show the changes in molecular weight of proteins (ovalbumin (Figure 3.6) and hCG (Figure 3.7)) after deglycosylation, as well tryptic mapping was used to identify any decrease in m/z of the protein following deglycosylation. The molecular weight of peptides produced following tryptic digestion was compared to the molecular weight of peptides available in the Mascot database, as described earlier (Chapter 2). Also, N-linked glycans released in this section were further identified and characterized.

Method

Ovalbumin or hCG (C0434) (10μg dissolved in 10μL of 50mM AMBIC), pH 8, was incubated at 100 °C for 10 minutes, before the addition of DTT (10μL of 10mM in AMBIC) and incubation at 56°C for 15 minutes. Next, IAA (10μL of 20mM in AMBIC) was added and tubes were incubated at room temperature for the following 15 minutes prior to the addition of 10μL (NBS, 1 activity units/μL) or 6 μL of PNGase F (Roche, 1.66 activity unit/μL). Thus, equivalent volumes corresponding to the same activity state were added respectively. Samples were gently vortexed and incubated at 37°C overnight. The sample was then loaded and aspirated 10 times onto solid-phase extraction tips (HyperCarb Porous Graphite Carbon (PGC) which allowed hydrophilic, polar glycans to bind. Next, to elute glycans from PGC tips 0.1% TFA in 50% ACN was used. Then, the eluent was dried in a vacuum centrifuge (30°C for 6h) reconstituted in 2μL of UHQ water, mixed with 10mg/mL HCCA/DHB matrix and spotted (1:1) on the MALDI target plate. MALDI-TOF MS parameters for protein identification comprised of positive linear mode with detector gain (63 x 3130V), m/z rage of 700 Da-60000

Da, reflector positive mode with m/z range of 700-5000Da, and the 6_ultra smartbeam parameter set with the zoom range of ±5%. Moreover, Mascot search parameters included SwissProt database, NCBI BLAST taxonomy browser limited to Homo sapiens species. Protein view: CGB7_Human, Choriogonadotropin subunit beta 7 OS=Homo sapiens OX=9606 GN=CGB7 PE=2SV=1. Experimental mass values were set to monoisotopic and variable modifications such as carbamidomethyl (C) and oxidation (M). The error window on the experimental peptide mass value was set to 500ppm. Mass values corresponded to charge carrier [M+H]+. The number of missed trypsin digest cleavages was set to 1, which is optimal for founding the target protein against the experimental data.

Results

The effect of PNGase F on ovalbumin from the two providers is shown in Figure 3.4. The effect of PNGase was established here by the shift of the m/z protein peak. The highest shift in m/z was detected after incubation with Roche PNGase (Figure 3.4 (d and e)), where the glycoprotein mass shifted from 44.6 000 Da to around 43 000 Da postdeglycosylation. Incubation with NBS PNGase, on the other hand, resulted in a significantly smaller peak shift (Figure 3.4 (b and c)). These results also suggest that PNGase F was more efficient when the glycoprotein was first denatured and reduced by a DTT as a peak shift of around 3000 Da was observed (Figure 3.4 (c and e)). Additionally, results of the deglycosylation process using Roche PNGase F were presented in Figure 3.5. Following incubation with PNGase (Roche), the corresponding decrease of the m/z ratio after deglycosylation was detected for $[M+H]^+$ and $[2M+H]^+$. The peaks were also seen on the chromatogram below (Figure 3.4), shifted to the left, corresponding to m/z reduction of around 3 000 Da after deglycosylation. As a result, DTT denaturation followed by Roche PNGase was used in all subsequent experiments.
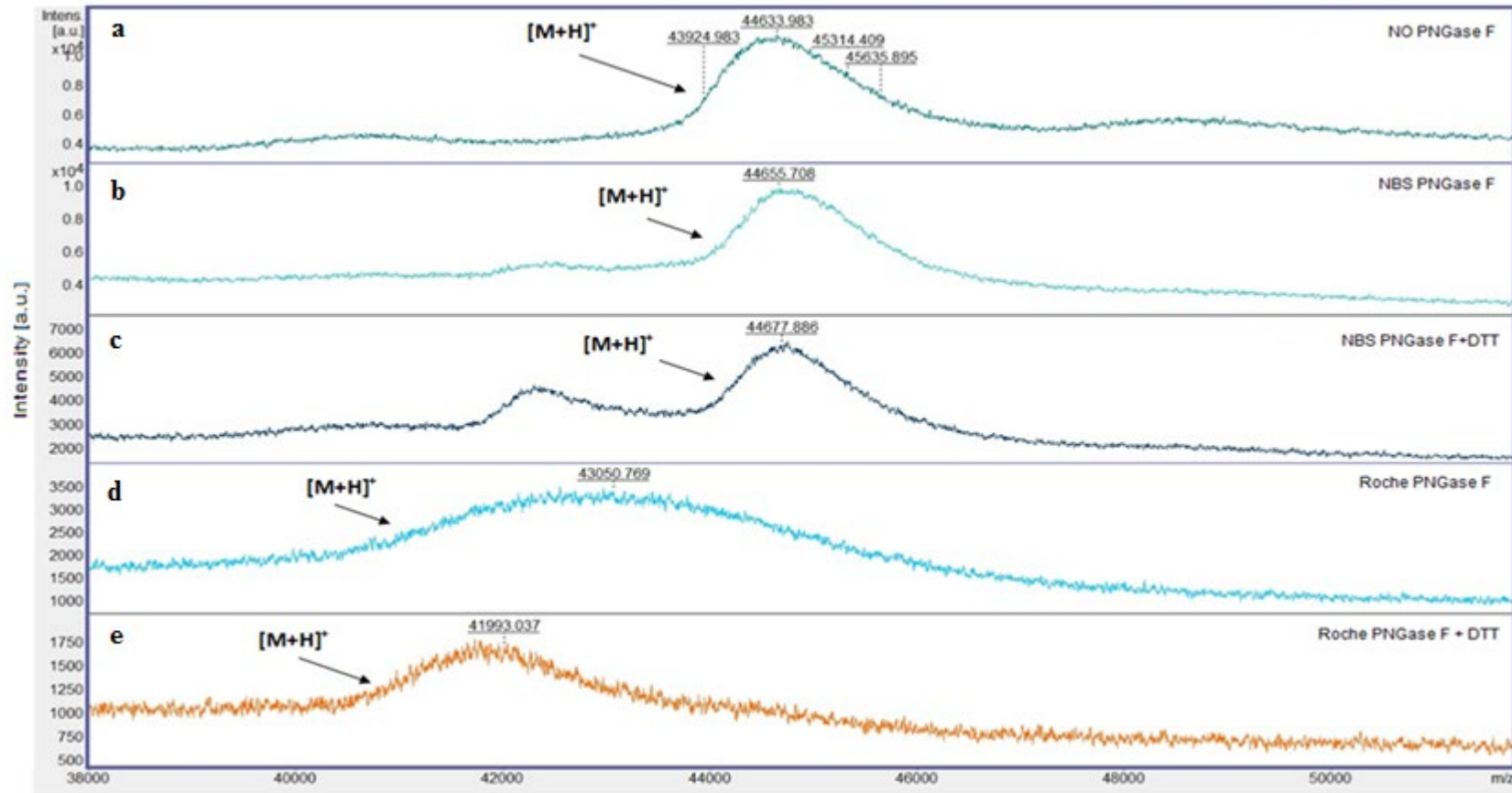
Figure 3.4 Magnified region of the ovalbumin peak [M+H]+ (44.600Da) visualized using MALDI-TOF after deglycosylation using NBS and Roche PNGases F. **a.** NO PNGase F, **b.** NBS PNGase F, **c.** NBS PNGase F + DTT, **d.** Roche PNGase F and **e.** Roche PNGase F + DTT
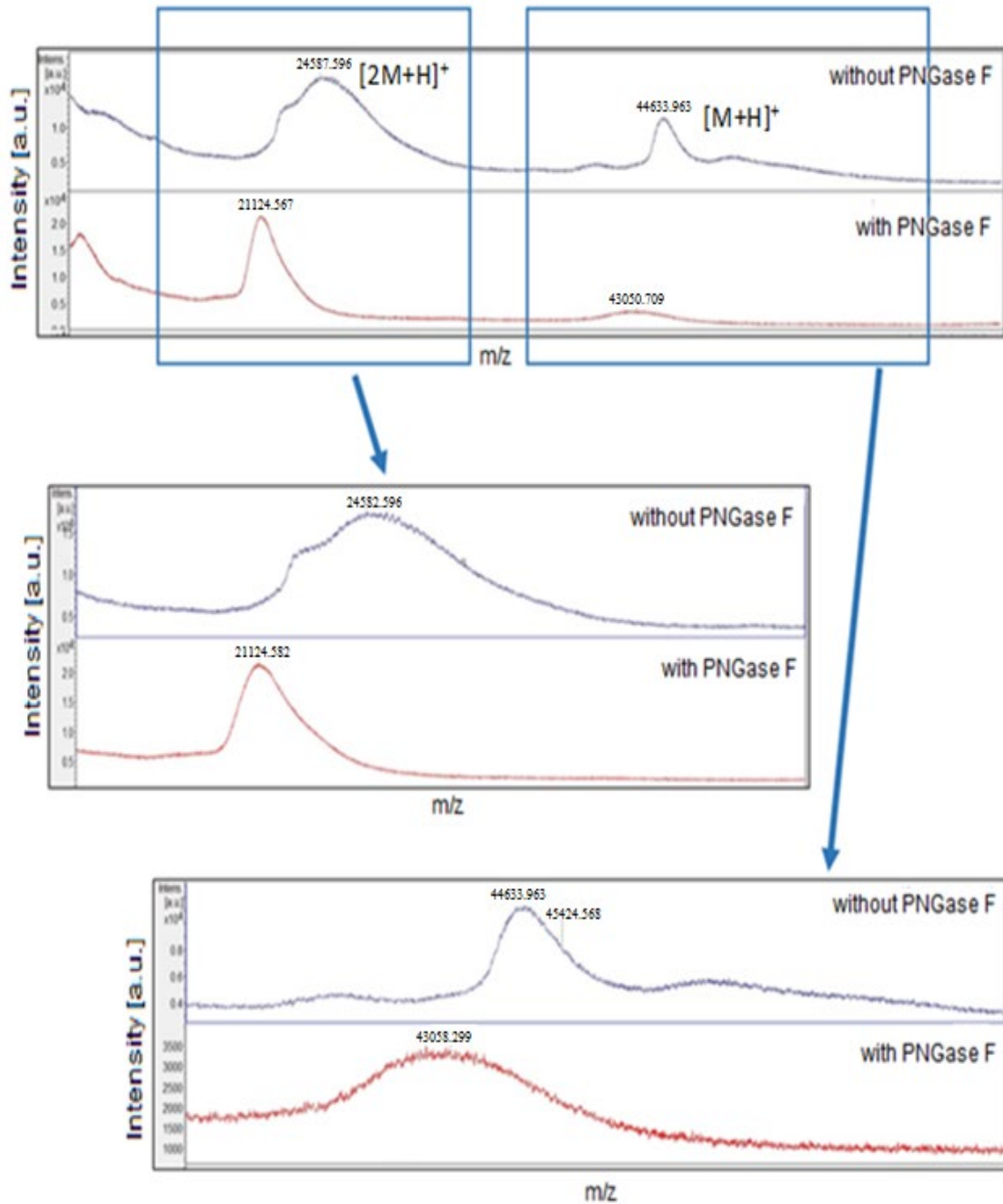
Figure 3.5 Effect of PNGase F (Roche, UK) on ovalbumin (44.600 Da). Zoom into peaks shows clearly the shift of m/z peaks which is caused by the removal of glycans (deglycosylation) using PNGase F and observed for [M+H]$^+$ and [2M+H]$^+$.

An optimized method for deglycosylation was then applied to hCG glycoprotein purified from human urine (C0434, Sigma, UK). The MALDI-TOF and SDS-PAGE were used to confirm changes in molecular weight of hCG after PNGAse treatment as presented in Figure 3.6 (a and b), respectively. The shift in the hCG protein peak mass (m/z) from 38 and 23 000Da (α + β and β chains respectively) to 35 and 19 000Da was observed after incubation with DTT, IAA and PNGase F at 37°C, overnight. This mass shift was confirmed following SDS-PAGE and Coomassie Blue staining of hCG samples with and without PNGase incubation (Figure 3.6 b). Results from identified released glycans found on hCG (C0434) are presented in the following Section 3.3.4.

a.
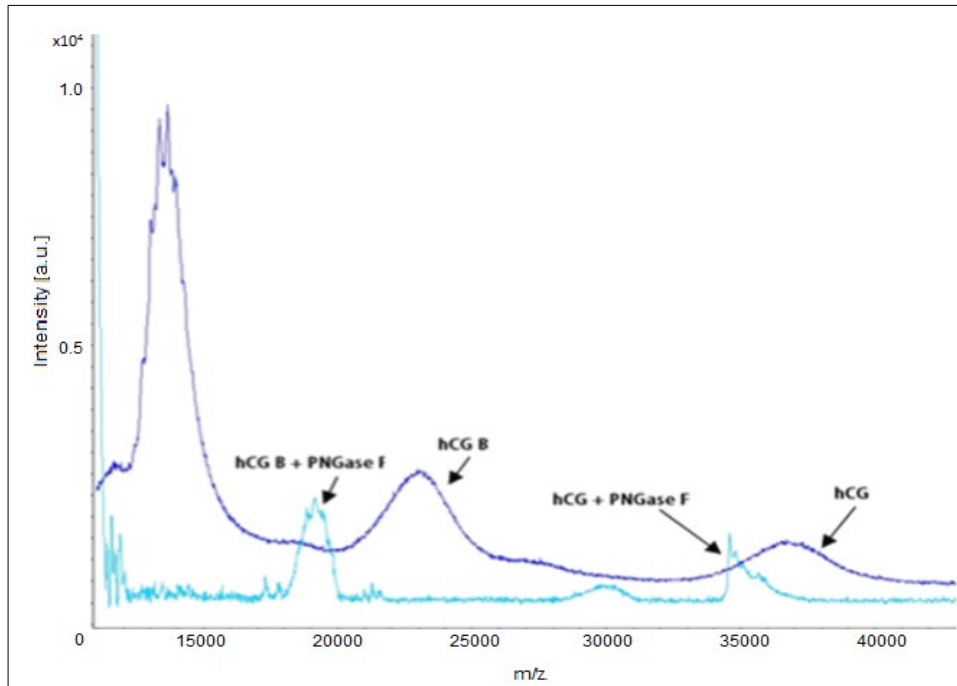
b.
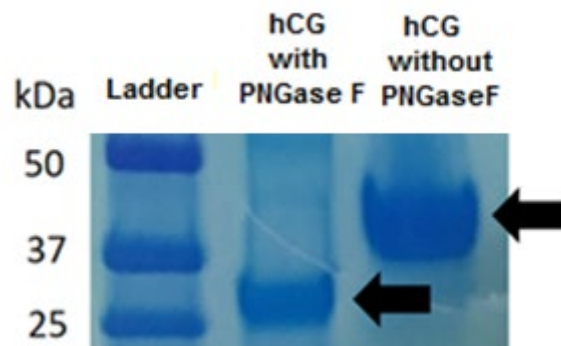
Figure 3.6 The effect of PNGase F on the intact hCG (C0434, Sigma) glycoprotein. **a.** The dark blue trace represents the hCG chromatogram using MALDI-TOF of hCG without PNGase (hCG peaks at approximately 38 kDa and 23 kDa). The light blue trace represents the hCG chromatogram in the presence of PNGase, hCG peaks at 35 kDa and 19 kDa can be seen. **b.** SDS-PAGE gel demonstrating the effect of PNGase F on the hCG. 38 kDa band for hCG and 35 kDa for treated with PNGase hCG can be seen (labelled with black arrows).

3.3 Methods for extraction of released N-linked glycans

The section above describes how PNGase F effectively released N-linked glycans from glycoproteins such as ovalbumin and hCG. This deglycosylation process resulted in a mixture of polar and nonpolar compounds such as proteins/peptides, as well as unlabelled and underivatized glycans, being detected in the sample. These proteins/peptides with hydrophobic properties might interfere with the ionisation of hydrophilic glycans, making MALDI-TOF MS analysis very challenging (Yang and Zhang, 2012). Therefore, solid-phase extraction (SPE) cartridges such as porous graphite carbon (Hypercarb, PGC), Carbon 18 (HyperSep C18), and Hydrophilic Interaction Chromatography (HILIC) tips were tested to purify highly polar glycans for further investigation. The porous graphite carbon (PGC) stationary phase tips can well retain and separate polar and hydrophilic molecules, such as N-linked glycans. This retention is due to the ability of molecules containing lone-pair or aromatic-ring electrons to interact through an electron transfer mechanism to the electronic cloud of the graphite (Pereira, 2008). Thus, highly polar compounds (N-linked glycans) can be enriched on a graphitized carbon column and easily eluted using low organic strength solvents (50% acetonitrile) (Yang and Zhang, 2012, Vreeker and Wuhrer, 2017). Further studies were undertaken to increase glycans extraction efficiency by increasing tip retention, analyte hydrophobicity and altering the interaction of glycans with the graphite surface by altering the nature of the organic solvent. Conversely, reverse-phase C18 resin stationary phase was used to retain highly non-polar compounds with a long-chain carboxylic acid such as proteins (Lochmüller *et al*., 1993) in order to separate them from polar glycans in the wash solution. Lastly, hydrophilic interaction chromatography (HILIC) tip separation is based on differences in hydrophilicity and hydrophobicity of glycans and peptides (Takegawa, *et al*., 2006).

3.3.1 Extraction of glycans from a sample mixture

The efficiency of glycan isolation from glycoprotein mixtures was compared using three types of solid-phase extraction tips i.e. HyperCarb Porous Graphite Carbon (PGC) tips, HyperSep Carbon18 tips and Hydrophilic Interaction Chromatography (HILIC) tips. Note that deglycosylation optimisation (Section 3.2) was investigated at the same time as the extraction of glycans from solid-phase extraction (SPE) (Section 3.3). As a consequence, in this part, we report the findings of the porous graphite carbon Hypercarb, PGC tips technique, which yielded the maximum number of N-glycans from glycoprotein to this point.

Method

Firstly 10µg of ovalbumin was dissolved in 10µL of 50mM AMBIC and heated (100 °C for 10 minutes) prior to the addition of 10µl of 10 mM DTT. The mixture was incubated at 56°C for 15 min) before the addition of 10µL of 20 mM IAA and 15 minutes incubation at room temperature. Then, 6µL of PNGase F, Roche (1.66 activity unit/µl) was added to a sample and incubated at 37°C overnight. The total volume of solution containing the detached glycan and protein mixture prepared during the deglycosylation procedure was 36µL. The mixture was then vortex mixed for 1min and loaded onto solid-phase extraction (SPE) tips i.e. C18 (Section 3.3.1.1), HILIC (Section 3.3.1.2) and HyperCarb (Section 3.3.1.3) described below. Finally, samples were then dried (30°C for up to 3h), reconstituted in 4µL of UHQ water, mixed with 10mg/mL HCCA/DHB matrix and spotted (1:1) on the target MALDI target plate. MALDI-TOF MS was set in a positive, reflector mode (m/z range of 700-5000 Da) and glycans were characterised using the GlycoQuest software.

3.3.1.1 Extraction using C18 tips

C18 Tips (Pierce®) were used for protein extraction. The protocol was adapted from the procedure recommended by the manufacturer. Extraction tips were equilibrated by aspirating 1 x 10µl of 50% ACN with 0.1%TFA, following 1x 10µl 0.1% TFA in water and then discarding the solvent. Aspirate up to 36 µl of deglycosylated 10µg ovalbumin (see Section 3.3.1 for full procedure) into the C18 tip and dispense and aspire (flushed 10 times) in order to retain all the non-polar molecules (deglycosylated protein). As glycans are very polar molecules they all should remain in the solution. This solution was then dried in a vacuum centrifuge (30°C for 5h) and reconstituted in 4µl of UHQ water, mixed with HCCA/DHB matrix and spotted (1:1) on the MALDI target plate. MALDI-TOF MS was set in a positive, reflector mode (m/z range of 700-5000 Da) and glycans were characterised using the GlycoQuest software.

3.3.1.2 Extraction using HILIC tips

The HILIC tip was equilibrated by aspirating 5 x 10µl of 15mM ammonium acetate/85% ACN, pH 3.5. 36 µl of glycan: ovalbumin mixture was loaded and aspirated 50 times to allow attachment of glycans to the tip. The tip was then washed 10x 10µl 50% ACN. 10 volumes of 10 µl 100% ACN was loaded and dispensed to elute the glycans. The resulting 100µl eluate was dried in a rotary vacuum evaporator for 4h at 30°C, reconstituted in 4µl of UHQ water, mixed with HCCA matrix and spotted (1:1) on the MALDI target plate. MALDI-TOF MS was

set in a positive, reflector mode (m/z range of 700-5000 Da) and glycans were characterised using the GlycoQuest software.
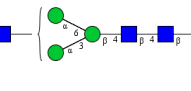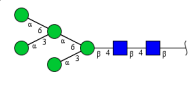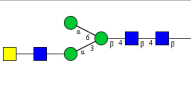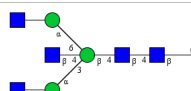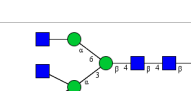
3.3.1.3 Extraction using PGC tips

The PGC tip was equilibrated by flushing twice with 10µl of methanol, following 4x 10 µL with 90 % can with 0.1% TFA and 5x 10µL 0.1% TFA and discarding the solvent in between each step. As above, for the maximum efficiency, 36 µl glycans mixture detached from ovalbumin (10µg) were slowly loaded and aspired 30 times. The tip was then flushed with 10x 10µl 0.1% TFA to remove salts from the tip. The glycans were then eluted by adding 20 volumes of 10 µl at a range of acetonitrile concentrations at 70% to elute the glycans. The resulting 200µl eluates were dried in a rotary vacuum evaporator for 6h at 30°C, reconstituted in 4µl of UHQ water, mixed with 10mg/mL HCCA/DHB matrix and spotted (1:1) on the MALDI target plate. MALDI-TOF MS was set in a positive, reflector mode (m/z range of 700-5000 Da) and glycans were characterised using the GlycoQuest software.
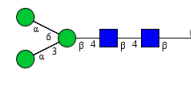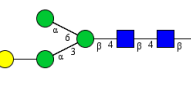
Results

Following extraction with C18, HILIC and PGC tips, glycans were extracted and detected using MALDI GlycoQuest (Table 3.4). While using C18 tips only a single Hex3HexNAc2 (m/z 933) glycan was identified. This was expected as C18 tips are designed to fast and efficiently capture non-polar protein and peptides and the technique was not precise enough to obtain hydrophilic molecules i.e. glycans from the solution containing salts and other possible polar contaminants. Next, extraction using the HILIC stationary phase, based on its ability to retain hydrophilic compounds, yielded two glycans Hex3HexNAc2 (m/z 933) and Hex4HexNAc2 (1095) however, extraction using porous graphite PGC tips yielded seven glycans as shown in Table 3.4.

Table 3.4 Comparison of solid-phase extraction tips glycan extraction from ovalbumin with HyperCarb Porous Graphite Carbon (PGC) tips, HyperSep Carbon18 tips and Hydrophilic Interaction Chromatography tips. ACN: acetonitrile

Glycans Structure	Glycan m/z & Composition	PGC 70% ACN (with 0.1%TFA)	HILIC 100% ACN	C18
	933 Hex3HexNAc2	+	+	+
	1095 Hex4HexNAc2	+	+	
	1136 Hex3HexNAc3	+		
	1257 Hex5HexNAc2	+		
	1339 Hex3HexNAc4	+		
	1542 Hex3HexNAc5	+		
	1663 Hex5HexNAc4	+		
	1745 Hex3HexNAc6			
	1866 Hex5HexNAc5			
	1948 Hex3HexNAc7			
	2110 Hex4HexNAc7			
	2151 Hex3HexNAc8			

3.3.2 Further optimisation of PGC extraction

In order to maximise the number of glycan species detected, the protocol for PGC extraction was repeated using further samples of ovalbumin following PNGase incubation to detach N-linked glycans. In this case, we used different ACN concentrations of 50, 70, 90 and 100% and resulting 200μl of eluants were dried in a rotary vacuum evaporator for 6h at 30°C, reconstituted in 4μl of UHQ water, mixed with 10mg/mL HCCA/DHB matrix and spotted (1:1) on the MALDI target plate. MALDI-TOF MS was set in a positive, reflector mode (m/z range of 700-5000 Da) and glycans were characterised using the GlycoQuest software.

Results

As before, 7 glycan species were seen following extraction of glycans from PGC tips using 70% ACN. These glycans were also detected in extracts using a lower proportion of ACN in the elutions mixture. However, here, additional higher molecular weight species were also detected. The optimum elution mixture contained 50% ACN with 0.1%TFA. Under these conditions, a total of twelve glycans were detected following MALDI-MS separation and GlycoQuest detection (Table 3.5).

We now show that the 50 % acetonitrile concentration utilised in the preceding sections for glycan elution is the best for this SPE approach, yielding the highest number of detected glycans. It was observed that increased acetonitrile concentrations cause ion suppression effect and reduced signal intensity during mass spectrometric detection. Thus, the number of N-linked glycans detected, decreased when the acetonitrile concentration used for elution was altered. From twelve glycans detected in 50% to seven in 70%, six in 90% and five in 100% with 0.1%TFA, as shown in Table 3.5. As a result of PGC's unique properties, highly polar glycans were easily eluted using lower organic strength solvents, and the analyte was retained as polarity increased. In conclusion, effective glycans detection and recovery was achieved by solid-phase extraction porous graphite carbon tips which were used in all subsequent experiments.

Table 3.5 Extraction of glycans from ovalbumin using PCG tips with eluent at 50, 70, 90 and 100% of acetonitrile. The composition of twelve detected glycans corresponds to N-linked glycans from ovalbumin following treatment with PNGase F (Roche) and detection using MALDI-TOF.

Glycans Structure	Glycan m/z & Composition	Acetonitrile concentration (%) PGC tips (with 0.1%TFA)			
		50%	70%	90%	100%
	933 Hex3HexNAc2	+	+	+	+
	1095 Hex4HexNAc2	+	+	+	+
	1136 Hex3HexNAc3	+	+	+	+
	1257 Hex5HexNAc2	+	+	+	+
	1339 Hex3HexNAc4	+	+	+	+
	1542 Hex3HexNAc5	+	+	+	
	1663 Hex5HexNAc4	+	+		
	1745 Hex3HexNAc6	+			
	1866 Hex5HexNAc5	+			
	1948 Hex3HexNAc7	+			
	2110 Hex4HexNAc7	+			
	2151 Hex3HexNAc8	+			

3.4 Detection – MALDI –TOF MS

So far in this chapter, the denaturing conditions (temperature, DTT and iodoacetamide concentrations) have been optimised for PNGase deglycosylation of ovalbumin. Three methods of solid-phase glycan extraction have been compared along with the further refinement of elution conditions for the preferred methods using PGC tips. In the following section further optimisation of the conditions for the separation and detection of glycans by MALDI –TOF MS are examined –namely instrument detection limits, the effect of a technique used for the sample application on the target plate and glycan concentration to peak intensity observed using MALDI-TOF.

Method

Mass spectrometric analysis was carried out using AutoFlex MALDI quadrupole ion trap-TOF instrument (Bruker, UK). Ionization was performed with a 337nm N_2 UV laser. Parent ion masses were measured in the reflectron mode with an accelerating voltage of 28 x 2106V, baseline 1.0% - 51.5 mV, 6_ultra smartbeam parameter set with the zoom range of ±5% and resolution set to 4.00GS. The total of the ion signals created by irradiating the target with 92% laser power was used to generate mass spectra, which were then calibrated internally using a prepared peptides calibrant Mix II from Bruker Daltonics, UK containing Bradykinin (757.40 Da), Angiotensin II (1046.54 Da), Angiotensin I (1296.68 Da), Substance P (1347.74 Da), Bombesin (1619.82 Da), ACTH (1-17) (2093.09 Da), ACTH (18-39) (2465.20 Da) and Somatostatin 28 (3147.47 Da) to offer a typical mass accuracy of 500 ppm with low or off real-time smoothing. Synthetic glycan standards NGAF2 (Ludger, CN-NGA2F) (Hex3HexNAc4Fuc1) of 1485.53 Da [M+Na] and glycan 1663.60 Da (Hex5HexNAc4) purified from human urine hCG molecule reconstituted in ultra-pure water were used for calibration and to study a limit of detection (Figure 3.7).

The detection mass range for glycans was set 800-4000 Da. The GlyoQuest database search method, version 1.3, was used to analyze glycan data as part of the Bruker ProteinScape database programme (version 3.0). The GlycoQuest matched m/z MS or MS/MS data against theoretical spectra calculated from databases entries. It worked by importing MS data (glycan pool spectra) and selecting a range of search parameters. Among others, type of glycan (hierarchical classification as defined in GlycomeDB database). The repertoire of N-linked glycans was obtained from public databases synchronized with ProteinDB, as well as a customised internal database (Search method: MG Glycan MALDI MS AA Carb, Method name: MG Glycan MALDI MS AA CarbBank) developed in Proteinscape and utilised for

GlycoQuest (glycan) searches. Glycan composition (only glycan structures that included the chosen forms/structures of monosaccharides were considered in the search), i.e. only N-Glycan core basic (minimum 3dXex+Hex2NAc). Modifications (e.g. sulphate, phosphate) include Sulphate (0-4), N-Sulphate (0-1), derivatization type (various labels) yet, in this study all glycans were underivatized. Next, min/max charges states were expected in the spectra, with multiple charges considered for this study. The ions (neutral) attached (ions defined as adducts providing a charge to the ions detected) i.e. H^+(0-1), Na^+(0-2), K^+(0-1), Li^+(0-1). These adducts and/or exchange ions that do not introduce a charge. The above parameter settings were used in a search that included neutral exchanges H+ vs. Na+, which were common in acidic glycans (exchange of acidic proton in sialic acid by Na+). Finally, the MS tolerance (m/z) detection was set to (m/z) 0.7 Da and monoisotopic peaks were detected. Glycan samples were mixed (1:1) with matrix mix of 10mg/mL 2,5-Dihydroxybenzoic acid (DHB) and 10mg/mL Alpha-Cyano-4-hydroxycinnamic acid (HCCA) in 50mM AMBIC/0.1% TFA and spotted on the MTP 394 ground steel ID210 1000476 target plate which was used throughout the study
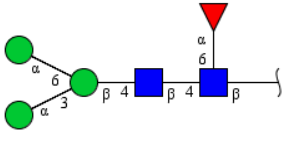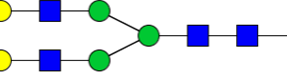
Glycan Structure	Glycan Composition	m/z [+Na adduct]
	Hex3HexNAc2dHex1	1485.53
	Hex5HexNAc4	1663.60
■ - N-acetylglucosamine ▼ - Fucose ● - Mannose		

Figure 3.7 Two biantennary N-linked glycans that contain terminal N-acetylglucosamine residues with Na+ adduct used for optimisation study. NGA2F (1485.53) also contain fucosylation on the core GlcNAc.

3.4.1 Instrument detection limit

The limit of blank (LOB) was calculated by assessing replicates of a blank sample and obtaining the mean and standard deviations of peak intensities. Replicates of a blank sample (n=20) were analysed and the mean value and standard deviation were derived to calculate the limit of detection (LOD), which was the lowest analyte concentration expected to be reliably distinguished from the blank (Armbruster and Pry, 2008). The LOD was calculated as a mean of blank intesity+ 3SD. As a result, the signal response on MALDI-TOF to a given amount of glycan was standardised to the amount of glycan giving rise to the response.

Method

Multiple, duplicate standards between 0.01 to 2.5 µg/µl of NGA2F (1485.53 Da) and hCG 1663.60 glycans were mixed 1:1 with 10mg/µl 2,5-Dihydroxybenzoic acid (DHB) and Alpha-Cyano-4-hydroxycinnamic acid (HCCA) in 50mM AMBIC/0.1% TFA) and spotted on the MALDI target plate to assess the sensitivity of the instrument. Glycans were identified using the parameters described above in Section 3.4.

Results

Two glycans 1485.53 Da (NGA2F) and hCG 1663.6 Da are shown in Figure 3.8 below. The glycan concentration between 0.01-2.5 µg produced a signal that was distinguishable from the signal-to-noise ratio (S/N) measured on the blank. Coefficients of variation (%CV) were determined for experimental replicates for each m/z peak based on the normalized intensity of the peak. For 1485. 342 and 1663.737 % CV equals 0.05 and 0.02, respectively. Thus, the limit of detection, defined as the lowest concentration that produced a signal intensity above 3SD from the mean of blank, was <20 ng/ml.
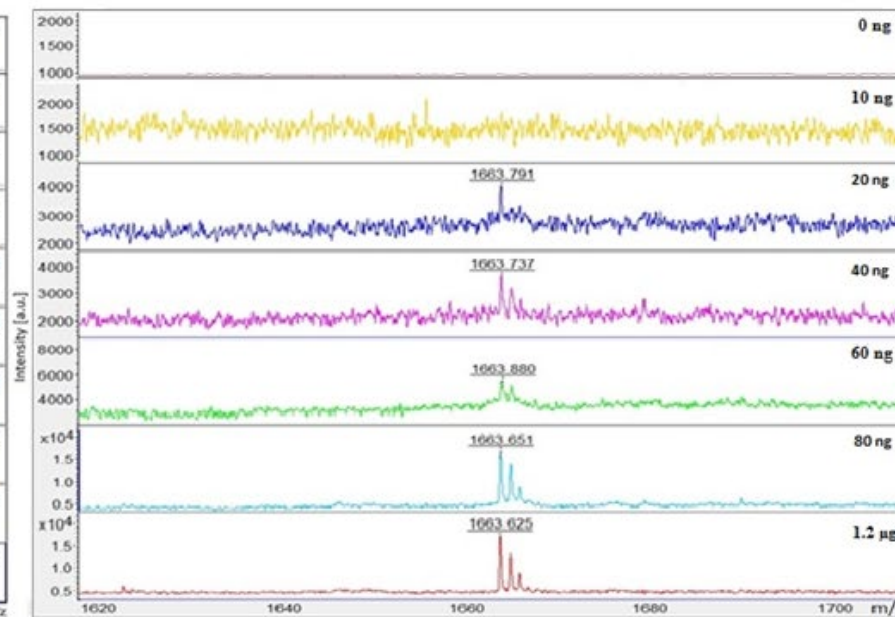
Figure 3.8 Chromatograms show monoisotopic peaks of N-linked glycans using MALDI-TOF MS. **a.** The NGA2F glycan with the theoretical mass of 1485.53 Da [M+Na] and **b.** glycan purified from hCG 1663.6 Da [Hex5HexNAc4].

3.4.2 Effect of different methods for matrix /sample application
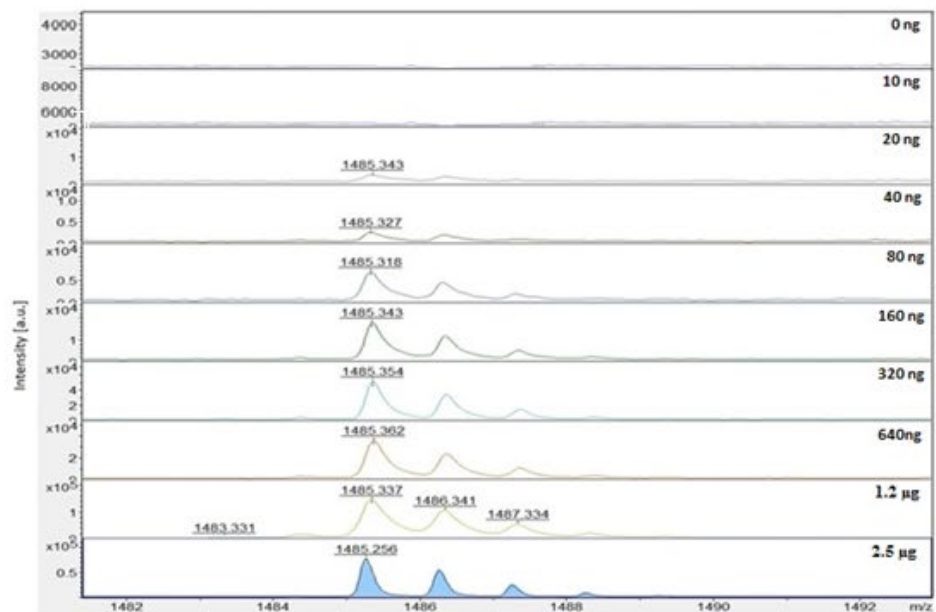
After the sample preparation and co-crystalization with matrix, commonly observed heterogeneous crystals might cause signal fluctuation after laser shot onto the sample. Therefore, three sample handling methods were evaluated to work on the homogeneity of sample distribution throughout the sample crystals. This side-by-side examination of each method was undertaken to maintain the reliability of all glycan analyses.

Method

1 µl of synthetic glycan standard NGAF2 (80ng/µL in H_2O) from Ludger, UK (Hex3HexNAc4Fuc1) was mixed with organic matrix mix (1:1) of 10mg/mL 2,5-dihydroxy benzoic acid and α-Cyano-4-hydroxycinnamic acid (DHB: HCCA) each was used for the detection of released N-linked glycans in positive ion mode using MALDI -TOF. As in the previous sections, matrices were reconstituted in 50% acetonitrile in 0.1% Trifluoroacetic acid 50% ACN/0.1% TFA to improve ionization for peptides and/or glycans separations using mass spectrometric analytical techniques. Multiple samples of glycans reconstituted in 2µL of UHQ water, mixed with 10mg/mL HCCA/DHB matrix and spotted (1:1) on the MALDI target plate in several combinations described in three methods below, were left to dry. MALDI-TOF MS reflector mode (m/z range of 700 Da-5000 Da) was used to detect glycans.

Method 1: Sample mixed with matrix and dried

1µL sample was mixed with 1 µL matrix and the mixture was dried at room temperature for 15 min then 1 µL was spotted onto the MALDI plate.

Method 2: Sample dried, matrix added and dried

1µL sample was spotted onto the MALDI plate and air-dried for 15 minutes. Next, 1 µL matrix was added onto the plate and left to air dry for additional 15 minutes at room temperature.

Method 3: Sandwich method: Matrix, sample, 2nd layer of matrix dried in between

1µL matrix air-dried was added onto the MALDI plate and left to air-dry for 15 minutes. Then, 1µL sample was added on top of the matrix and left to air dry for 15 minutes at room temperature. Finally, the second layer of the matrix was added and left to dry for 15 minutes at room temperature.

Results

It was found that there is a significant effect on the intensity of the MALDI-TOF signal obtained in relation to the used method for sample spotting. A 20% spike in peak intensity was seen where the sample was mixed with matrix and dried or where the matrix was dried prior to the addition of samples (method 1) However, the method 2 resulted in a 4 fold increase in the peak intensity (80%). Special attention needs to be applied to maintain homogeneity of sample distribution throughout the sample crystals. The differences between the three methods were presented in Figure 3.9 (n=5 for each method, ANOVA, $p<0.05$). Method 2 (sample dried, matrix added and dried) was applied in all further experiments.
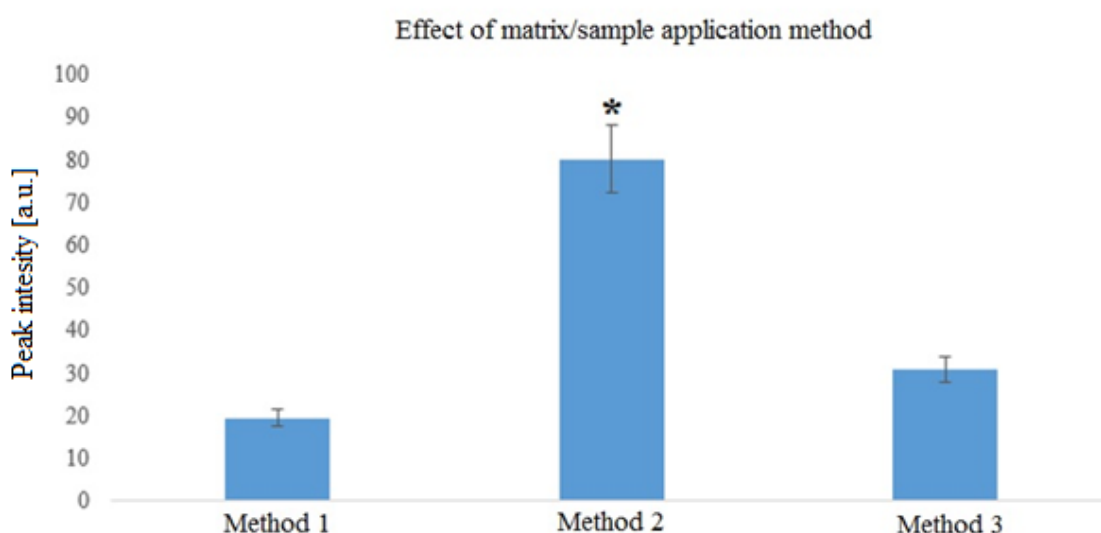

Figure 3.9 Comparison of three sample application methods onto target plate for MALDI-TOF analysis (n=5 for each method, * denotes a significant increase compared to other methods (ANOVA, $p<0.05$).

3.4.3 Effect of eluent concentration to peak intensity

After the sample preparation and co-crystalization with matrix using, method 2 described above (Section 3.4.2), homogenous crystals were obtained. To determine if the amount of glycan on each spot was optimal for accurate detection, the effect of glycan concentration (in elution and wash) on peak intensity was undertaken. Elution and wash of PGC tips loaded with various concentrations of commercially available synthetic glycan standard NGAF2 (Ludger, UK) were analysed using MALDI-TOF MS and GlycoQuest software.

Method

The PGC tip was equilibrated as described in Section 3.3.1.3. Briefly, the tip was emerged in 10µl of methanol x 2, following 4 x 10 µL with 90 % ACN/0.1% TFA and 5x 10µL 0.1% TFA and discarding the solvent in between each step. Next, different concentration (20-160 ng/µL) of 10µL NGAF2 (Ludger, UK) was slowly loaded and aspired 30 times. The tips were then flushed with 10x 10µl 0.1% TFA to remove salts (and wash solution retained for further analysis) and glycans were eluted with 20x10 µl of 50% acetonitrile with 0.1%TFA. The resulting 200µl eluates were dried in a rotary vacuum evaporator for 5h at 30°C and reconstituted in 4µl of UHQ water. 1µL of the sample was spotted onto the MALDI plate and air-dried for 15 minutes. Next, a 1 µL matrix was added onto the plate and left to air dry for additional 15 minutes at room temperature, each sample was repeated five times. MALDI-TOF MS was set in a positive, reflector mode (m/z range of 700-5000 Da) and glycans were characterised using the GlycoQuest software.

Results

The effect of glycan concentration on the peak intensity is shown in Figure 3.10. It was found that the peak intensity increased as the amount of glycan loaded was increased from 20 to 40 ng of glycan. However, as the glycan concentration was raised further to 80ng and 160ng, respectively, the peak intensity decreased. Thus the intensity of the peak was relatively connected to the amount of analysed glycan on the target plate. Hence, the optimal glycan concentration was 40ng/µL of glycan sample obtained from elution. Also, the relative intensity of eluent compared to background (wash) was presented in Figure 3.10.
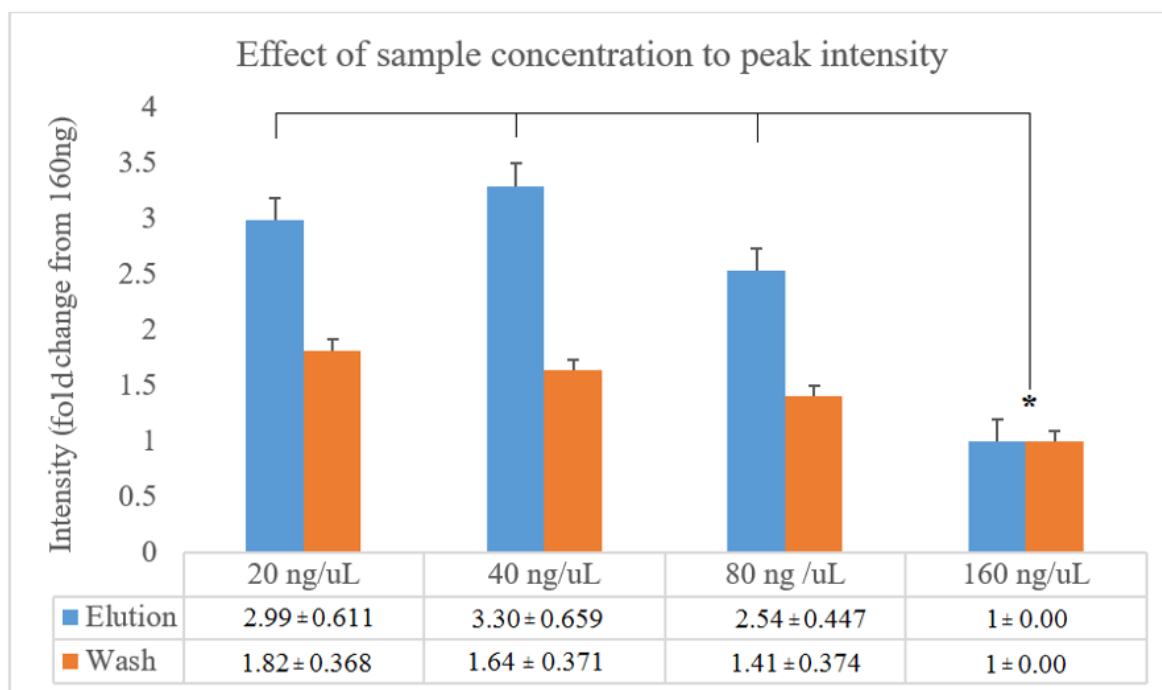
Figure 3.10 Relative peak intensity between elution and wash sample using MALDI-TOF. The relative intensity of eluent compared to background (wash) presents high effectiveness of detection at a concentration in the range of 20-80 ng of glycan. (mean n=5, mean $\pm$ SEM, * denotes significant difference from other three concentrations, ANOVA, $p<0.05$)

3.5 Comparison of N-linked glycan detection with Orbitrap-MS

In Section 3.4, we focused on enhancing the detection of the smallest amount of glycan in the sample, which could still be detected but not necessarily quantified under the specified experimental conditions. Furthermore, our results confirmed that the process of sample and matrix application, as well as the final concentration of the sample on the MADI-TOF target plate, have a substantial influence on the peak intensity (from standard glycan) that follows the identification of released N-linked glycans from the glycoprotein. As a consequence, the study of released N-linked glycans was continued using previously explained steps and applied to ovalbumin and hCG glycoproteins. In addition, the LC-MS/MS QExactive Orbitrap mass spectrometer was employed to identify N-linked glycan from hCG.

Method
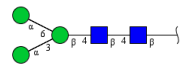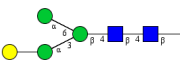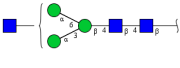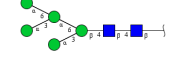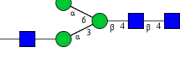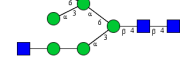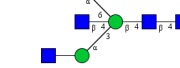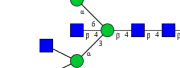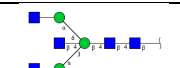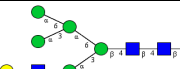
Ovalbumin (Sigma, UK) and hCG (Fitzgerald®, AC048 and Sigma, C0434) (10μg dissolved in 10μL of 50mM AMBIC), were incubated at 100 °C for 10 minutes before the addition of 10μl DTT of 10mM in AMBIC followed by incubation at 56°C for 15 minutes. Next, 10μL of 20 IAA in AMBIC was added and tubes were incubated at room temperature for 15 minutes. Next, 6μl PNGase F was added and left for overnight incubation at 37°C. For glycan purification PGC tips were used, equilibrated by flushing twice with 10μl of methanol, following 4x 10 μL with 90 % ACN/0.1% TFA and 5x 10μL 0.1% TFA and discarding the solvent in between each step. 36 μl glycans mixture detached from glycoproteins were slowly loaded and aspired 30 times onto a tip. Then flushed with 10x 10μl 0.1% TFA to remove salts and then eluted by adding 20 volumes of 10 μl of 50% to elute the glycans. The resulting 200μl eluates were dried in a rotary vacuum evaporator for 6h at 30°C and reconstituted in 4μl of UHQ water. The technique used for applying the sample onto the MALDI target plate included spotting 1μL of the sample following air-drying for up to 15 minutes, next adding 1 μL of 10mg/mL HCCA/DHB matrix on top and air-drying for additional 15 minutes at room temperature. Subsequently, MALDI-TOF MS was set in a positive, reflector mode (m/z range of 700-5000 Da) with an acceleration voltage of 28 x 2106V, baseline 1.0% - 51.5 mV, 6_ultra smartbeam parameter set with the zoom range of ±5% and resolution set to 4.00GS. Profiling of released glycans was performed using GlycoQuest software from Bruker, UK and GlycoWorkbench® available online. Also, at least three rounds of MALDI TOF data was collected for each sample and pooled results were analysed using GlycoQuest software. Furthermore, to compare obtained results using MALDI-TOF, LC-MS/MS QExactive Orbitrap mass spectrometer (ThermoFisher, UK) was used. Glycan samples were analysed using a

Dionex Ultimate 3000 RSLC Nano ultra-high performance liquid chromatography system coupled to a QExactive Orbitrap mass spectrometer. The instrument was reconfigured to enable direct injection of samples via a 6 port Rheodyne® injection valve fitted with a 5 µL loop. This ensured that any unretained molecules would flow into the electrospray source facilitating detection and analysis. Furthermore, the pre-column was re-positioned directly before the electrospray source to retain residual deglycosylated protein, hence affording a crude, but effective separation. The concentrated sample was separated using a binary gradient elution profile composed of a mixture of water and acetonitrile (95:5 v/v) containing 0.1 % formic acid (eluent A) and a mixture of acetonitrile and water (80:20 %, v/v) (eluent B) containing 0.1 % formic acid at a flow rate of 6 µLmin-1. The gradient was 0 min- 0 %B, 4 min5 %B, 5 min-8%B, 40 min-40 %B, 41 min-80%B. The autosampler and column oven temperature was set to 4 and 40°C, respectively. The QExactive was operated in a data-dependent mode. MS survey scans were acquired from m/z 400 to 2500 at a resolution of 35,000 with AGC of 3e6 and a maximum IT of 100 ms. The 10 most abundant ions were subjected to MS/MS and measured with a resolution of 17,500 and AGC of 1e5 and maximum IT of 200 ms.

Results

The previous method described in Section 3.3.1 allowed for the detection of twelve N-linked ovalbumin glycans, which are listed in Table 3.5; however, after further optimization described in Section 3.5, the detection was improved by 43 % per, resulting in the characterisation of 21 structures. Thus, additional nine distinct glycan compositions from ovalbumin (Sigma, UK) that coincide with those reported in the Sigma-Aldrich respiratory and Wilson, *et al.,* (2000) were characterised and presented in Table 3.6. As a result, Figure 3.11 presented the MALDI-TOF chromatogram of twenty-one N-linked glycan structures found on ovalbumin. Glycans ranged from m/z 933 Da to 2151 Da in size. They contained the common N-glycan core Man3GlcNAc2Asn and multiple hexoses and GlcNAc residues moreover, six bisected glycans structures were detected.

Table 3.6 Twenty-one individual glycan structures and compositions (1-21) of glycans released from ovalbumin. The values of peak intensity correspond to the intensity of peaks detected on the chromatogram (MALDI-TOF MS). Monoisotopic glycan masses were measured in the positive ion reflectron mode as [M+Na$^+$] adducts (Figure 3.11).

	Glycan Structure	Glycan m/z & composition
1.		**933** Hex3HexNAc2
2.		**1095** Hex4HexNAc2
3.		**1136** Hex3HexNAc3
4.		**1257** Hex5HexNAc2
5.		**1298** Hex54HexNAc3
6.		**1339** Hex3HexNAc4
7.		**1419** Hex6HexNAc2
8.		**1460** Hex6HexNAc3
9.		**1501** Hex4HexNAc4
10.		**1542** Hex3HexNAc4
11.		**1581** Hex3HexNAc5
12.		**1622** Hex6HexNAc3
13.		**1663** Hex5HexNAc4

14.		**1704** Hex5HexNAc5
15.		**1745** Hex3HexNAc6
16.		**1866** Hex5HexNAc5
17.		**1907** Hex4HexNAc6
18.		**1948** Hex3HexNAc7
19.		**2028** Hex6HexNAc5
20.		**2110** Hex4HexNAc7
21.		**2151** Hex3HexNAc8
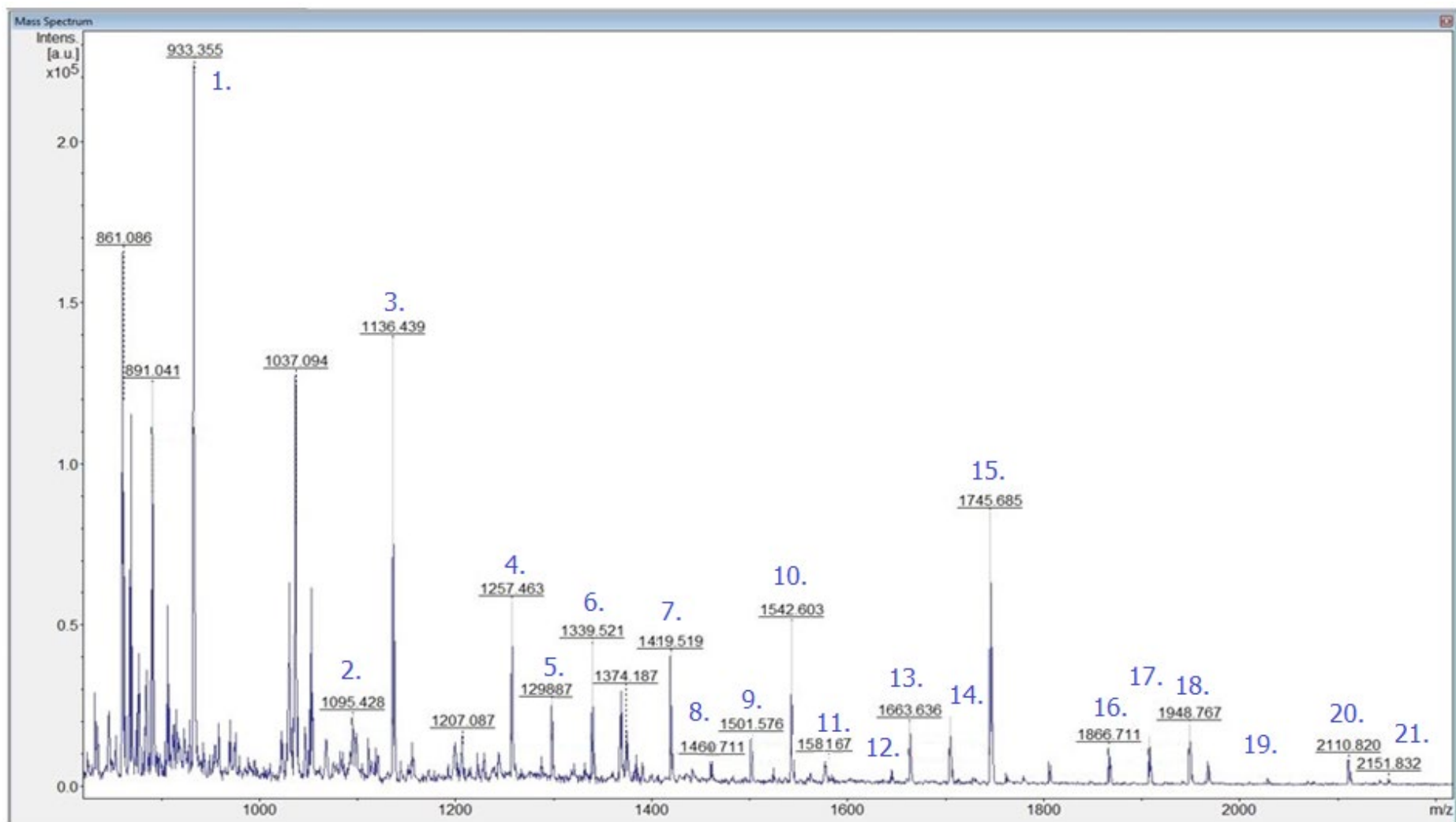
Figure 3.11 Representation of all twenty-one glycan structures found on ovalbumin, with an assigned number (1-21), m/z [M+Na+] for peak 12., 14., 17. and 19. is 1622 Da, 1704 Da, 1907 Da and 2028 Da, respectively using MALDI-TOF as indicated in Table 3.6.

A successful method for the release and detection of glycans from ovalbumin was later used with hCG standards isolated from human pregnant urine. Table 3.7 and Table 3.8 displayed fifty N-linked glycans isolated from human urine (different gestational weeks) from three commercially available intact hCG standards using MALDI-TOF MS and LC-MS Orbitrap. In detail, Table 3.7 represents glycan profiles released from three hCG standards: AC048 (Fitzgerald Laboratories, USA), C0434 Batch 1 and C0434 Batch 2 (Sigma, UK.). Each sample was analysed 4 times to maximise the number of glycans detected and the final data presented in Table 3.7 and 3.8 shows combined glycans structures from all of these samples. It is very important to mention that all presented structures are postulations only and complete compositional assignments require further experiments. We are fully aware of MALDI-TOF limitations, therefore the raw data accumulated using MALDI-TOF MS was deconvoluted into the peak list, which was matched with the library of theoretical spectra found in various glyco-informatic databases (i.e. CarbBank, Glycome DB, UniCarb KB) used by GlycoQuest software. Search parameters in GlycoQuest included N-glycan containing pentasaccharide core structures (Hex3GlcNAc2), with multiple ions adducts reported for MALDI-MS analysis and water ion adducts for LC-MS. As well, for more accurate speculation of compositional glycan assignments the theoretical m/z values, were created using GlycoWorkBench software (see Table 3.7). Thus, varying glycan distribution, amount, and composition using MALDI-TOF and LC-MS Orbitrap hypothesised glycans configuration across all samples, including core and multiple fucosylation, a substantial number of truncated glycans, and glycans with several sialic acid groups. In AC048 twenty fucosylated, sialylated, and biantennary glycans were detected and compared to sixteen related structures found by LC-MS Orbitrap. In addition, twenty glycans were identified, including multi-mannose in C0434 Batch1 and eight fucosylated, sialylated glycans.

Table 3.7 Representation of combined adducts + glycan complex released from three hCG standards: AC048 (Fitzgerald), C0434 Batch 1 and C0434 Batch 2 (Sigma). Each standard was analysed four times and combined data was presented. Composition and m/z values of glycans were detected using MALDI-TOF and LC-MS Orbitrap. A mixture of sodium, potassium and lithium adducts were reported for MALDI-MS analysis and water ions adducts for LC-MS. Theoretical m/z values were measured using GlycoWorkBench software. Hex-Hexose, HexNAc- N- Acetylglucosamine, Neu- Sialic Acid, dHex- Fucose
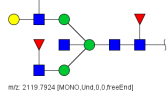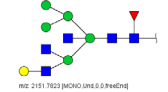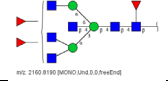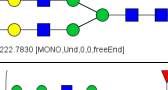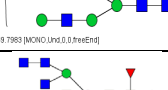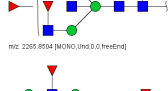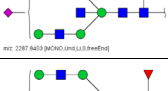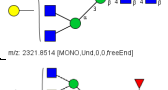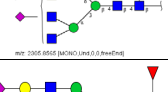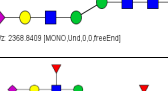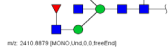
Theoretical Glycan m/z	Glycans composition from hCG	Assembly	Adduct	MALDI-TOF AC048 m/z	MALDI-TOF C0434 Batch 1 m/z	MALDI-TOF C0434 Batch 2 m/z	LC-MS Orbitrap AC048 m/z
MONOANTENNARY							
910.32	Hex3HexNAc2		H_2O^+	910.1		910.3	
910.32	Hex3HexNAc2		Na^+	933.3	933.3		
910.32	Hex3HexNAc2		K^+		948.7		
1072.38	Hex4HexNAc2		Na^+	1095.9			
1056.38	Hex3HexNAc2dHex1		K^+		1095.9		
1072.38	Hex4HexNAc2		$H+$		1111.3		
1113.40	Hex3HexNAc3		Na^+		1136.4		
1858.67	Hex4HexNAc3dHex2NeuAc1	Fucosylated Sialylated	H_2O^+				1860.01
BIANTENNARY							
1234.43	Hex5HexNAc2	High mannose	Na+	1257.4	1257.4		
1259.45	Hex3HexNAc3dHex1	Fucosylated	Li+			1266.8	
1275.46	Hex4HexNAc3	Bisected	Na+	1298.4	1298.4		
1316.48	Hex3HexNAc4	Bisected	K+		1345.7		
1234.43	Hex5HexNAc2	High mannose	Li+		1374.1		
1421.51	Hex4HexNAc3dHex1	Fucosylated	K+	1460.3	1460.4		
1519.56	Hex3HexNAc5	Bisected	Na+		1498.3		
1599.56	Hex6HexNAc3	High mannose	Na+	1622.6	1622.5		
1624.59	Hex4HexNAc4dHex1	Fucosylated	Na+	1663.4			
1712.61	Hex4HexNAc3NeuAc1dHex1	Fucosylated Sialylated	H^+	1713.3	1713.1		
1916.71	Hex4HexNAc4dHex3	Fucosylated	Na^+ H_2O^+	1939.1			1916.12
1898.67	Hex3HexNAc4NeuAc2	Sialylated	K^+		1937.5		
2425.86	Hex5HexNac5NeuAc2	Sialylated	H^+	2222.8			2222.10
2239.79	Hex6HexNAc4NeuAc1dHex1	Fucosylated Sialylated	H^+	2240.6			
2368.84	Hex5HexNAc4NeuAc2dHex1	Fucosylated Sialylated	K^+ H_2O^+	2389.1			2368.78
TRIANTENNARY							
1956.71	Hex3HexNAc5NeuAc1dHex1	Fucosylated Sialylated	Li^+	1957.7	1963.8		
1973.73	Hex4HexNAc5dHex2	Fucosylated	Li^+	1976.0			1973.47
1973.73	Hex4HexNAc5dHex2	Fucosylated	Na^+ H_2O	2099.5			2094.69
2119.79	Hex4HexNAc5dHex3	Fucosylated	K^+ H_2O^+	2141.2			2120.97
2222.78	Hex5HexNAc4NeuAc2	Sialylated	H^+	2264.9	2262.6		
2240.81	Hex6HexNAc4dHex3 Hex4HexNAc5NeuAc1dHex2	Fucosylated Sialylated	Na^+ H_2O^+	2271.0			2265.55
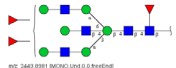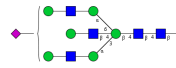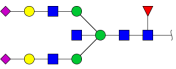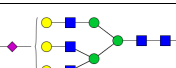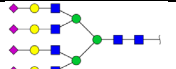
2240.81	Hex6HexNAc4dHex3	Fucosylated	K+	2279.3			
2281.84	Hex5HexNAc5dHex3	Fucosylated	H+	2282.7		2283.3	
2280.82	Hex5HexNAc5NeuAc1dHex1	Fucosylated Sialylated	Li+ H2O+	2288.2			2281.19
2410.88	Hex4HexNAc5NeuAc1dHex3	Fucosylated Sialylated	Na+ H2O+	2415.1			2410.03
2160.81	Hex3HexNac6dHex3	High mannose	Na+	2183.8	2183.6		
2443.89	Hex6HexNAc5dHex3	Fucosylated	H+			2445.4	
2499.89	Hex6HexNAc6NeuAc1	Sialylated	Li+			2507.3	
2571.92	Hex5HexNAc5NeuAc2dHex1	Fucosylated Sialylated	Na+ H2O+	2576.2			2571.91
2587.91	Hex6HexNAc5NeuAc2	Sialylated	K+ H2O+	2609.0			2587.98
2733.97	Hex6HexNAc5NeuAc2dHex1	Fucosylated Sialylated	Na+ H2O+	2737.1			2733.97
TETRANTENNARY							
1802.64	Hex6HexNAc4	High mannose	K+ H2O+	1841.6	1842.3		1802.98
1925.72	Hex3HexNAc7		H+		1927.2		
2151.78	Hex6HexNAc5dHex1	Fucosylated	Na+ H2O+	2173.3			2151.38
1957.73	Hex3HexNAc6dHex3	Fucosylated	Na+		2183.2		
2233.83	Hex4HexNAc7dHex1	Fucosylated	Na+	2256.3			
2612.94	Hex4HexNAc6NeuAc2dHex1	Fucosylated Sialylated	H2O+				2319.45
2305.85	Hex3HexNAc6NeuAc1dHex2	Fucosylated Sialylated	K+	2344.7			
2955.08	Hex7HexNAc6dHex4	Fucosylated	K+			2994.6	
3535.23	Hex7HexNAc6NeuAc4	Sialylated	Li+			3542.0	
3722.32	Hex6HexNAc7NeuAc4dHex1	Fucosylated Sialylated	H+			3758.1	

Table 3.8 Fifty glycan structures released from hCG standards detected using MALDI-TOF and LC-MS Orbitrap. Variation of mono-, bi-, tri- and tetra- antennary or high mannose branching with fucosylation, sialyation and/or truncated glycans were detected from three hCG standards: AC048 (Fitzgerald), C0434 Batch 1 and C0434 Batch 2 (Sigma). Multiple adducts M+Na+, M+K+, M+Li+ or water were detected.

	Glycan composition	Glycan structure
1	**Hex3HexNAc2**	m/z: 910.3278 [MONO,Und,0,0,freeEnd]
2	**Hex4HexNAc2**	m/z: 1095.3698 [MONO,Und,Na,0,freeEnd]
3	**Hex3HexNAc2dHex1**	m/z: 1095.3488 [MONO,Und,K,0,freeEnd]
4	**Hex4HexNAc2**	m/z: 1111.3438 [MONO,Und,K,0,freeEnd]

5	**Hex3HexNAc3**	m/z: 1136.3964 [MONO,Und,Na,0,freeEnd]
6	**Hex5HexNAc2**	m/z: 1257.4226 [MONO,Und,Na,0,freeEnd]
7	**Hex3HexNAc3dHex1**	m/z: 1266.4805 [MONO,Und,Li,0,freeEnd]
8	**Hex4HexNAc3**	m/z: 1298.4492 [MONO,Und,Na,0,freeEnd]
9	**Hex3HexNAc4**	m/z: 1355.4497 [MONO,Und,K,0,freeEnd]
10	**Hex4HexNAc3dHex2NeuAc1**	
11	**Hex4HexNAc3dHex1**	m/z: 1460.4810 [MONO,Und,K,0,freeEnd]
12	**Hex5HexNAc3**	m/z: 1498.4579 [MONO,Und,K,Na-H,freeEnd]
13	**Hex6HexNAc3**	m/z: 1622.5548 [MONO,Und,Na,0,freeEnd]
14	**Hex4HexNAc4dHex1**	m/z: 1663.5604 [MONO,Und,K,0,freeEnd]
15	**Hex4HexNAc3NeuAc1dHex1**	m/z: 1712.6133 [MONO,Und,0,0,freeEnd]
16	**Hex6HexNAc4** **(two possible structures)**	m/z: 1802.6450 [MONO,Und,0,0,freeEnd]
17		m/z: 1841.6082 [MONO,Und,K,0,freeEnd]
18	**Hex4HexNAc3NeuAc1dHex2**	m/z: 1858.6712 [MONO,Und,0,0,freeEnd]
19	**Hex4HexNAc4dHex3**	m/z: 1916.7131 [MONO,Und,0,0,freeEnd]
20	**Hex3HexNAc7**	m/z: 1925.7246 [MONO,Und,0,0,freeEnd]
21	**Hex3HexNAc4NeuAc2**	m/z: 1937.6405 [MONO,Und,K,0,freeEnd]
22	**Hex3HexNAc5NeuAc1dHex1**	m/z: 1963.7347 [MONO,Und,Li,0,freeEnd]
23	**Hex4HexNAc5dHex2**	m/z: 1973.7345 [MONO,Und,0,0,freeEnd]
24	**Hex6HexNAc4dHex2**	m/z: 2094.7608 [MONO,Und,0,0,freeEnd]

120

25	Hex4HexNAc5dHex3	
26	Hex6HexNAc5dHex1	
27	Hex3HexNac6dHex3	
28	Hex5HexNAc5NeuAc2	
29	Hex6HexNAc4NeuAc1dHex1	
30	Hex4HexNAc7dHex1	
31	Hex5HexNAc4NeuAc2	
32	Hex4HexNAc5NeuAc1dHex2	
33	Hex6HexNAc4dHex3	
34	Hex6HexNAc4dHex3	
35	Hex5HexNAc5dHex3	
36	Hex5HexNAc5NeuAc1dHex1	
37	Hex5HexNAc5NeuAc1dHex1	
38	Hex3HexNAc6NeuAc1dHex2	
39	Hex4HexNAc6NeuAc2dHex1	
40	Hex3HexNAc6NeuAc1dHex2	
41	Hex5HexNAc4NeuAc2dHex1	
42	Hex4HexNAc5NeuAc1dHex3	

121

43	Hex6HexNAc5dHex3	
44	Hex6HexNAc6NeuAc1	
45	Hex5HexNAc5NeuAc2dHex1	
46	Hex6HexNAc5NeuAc2	
47	Hex6HexNAc5NeuAc2dHex1	
48	Hex7HexNAc6dHex4	
49	Hex7HexNAc6NeuAc4	
50	Hex7HexNAc7NeuAc4dHex1	
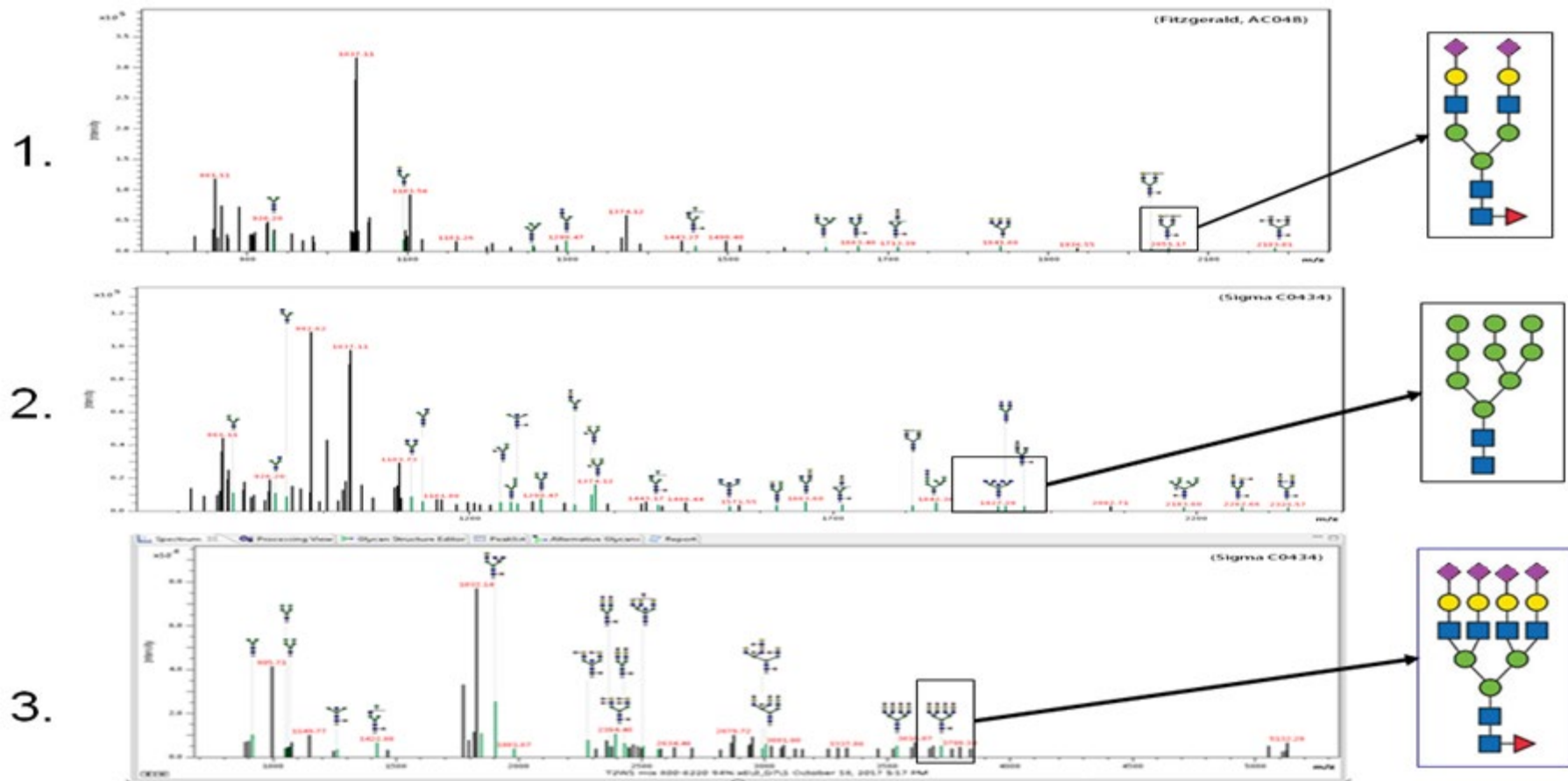
Figure 3.12 Mass spectrometry of N-linked glycans released from three different hCG standards. **1.** Fitzgerald (AC048), **2.** Sigma (C0434, Batch1), **3.** Sigma (C0434, Batch2) with assigned m/z and glycan composition. Variation of glycan composition from fucosylated, sialylated, biantennary in AC048, to multi-mannose in C0434, Batch1 and fucosylated, sialylated, tetrantennary in C0434, Batch2 were observed in all samples.

3.6 Discussion

In this chapter, we have optimised methods for deglycosylation, extraction and sample preparation for the detection and characterisation of N-linked glycans (Figure 3.1). Briefly, the deglycosylation process aimed to remove N-linked glycans by the enzymatic release, this was followed by purification of glycans using graphitized carbon chromatography tips. Thus, this methodology enables the detection of as little as 20ng of glycan and allows the detection of fifty glycans structures released from hCG.

Firstly, to release of N-linked glycans from hCG peptide-N-glycosidase F enzyme (PNGase F) was used (Figures 3.4 and 3.5). PNGase F breaks the amide link between asparagine and glycosyl amine, releasing the glycosyl amine, which may then be easily transformed into a glycan with a reducing end (aldehyde functional group) (Anumula, 2006). Thus, analysis of the complex mixtures of proteins/peptides and glycans by MALDI-TOF MS might be limited by the phenomenon of ion suppression, in which the most abundant molecules suppress ionization of those in lower abundance (Figure 3.10) (Zaia, 2008). To enhance the purity of released N-linked glycans and improve detection, the solid-phase extraction (SPE) applications of porous graphitic carbon (PGC) as a chromatographic stationary phase, were used (Section 3.3.1.3). This technique also become very useful in the separation of underivatized N-glycans as described by Jensen *et al*., 2012 and more recently by Young *et al.,* 2021, where released glycans were detected in positive mode.

Thus, enrichment with PGC tips was ideal for hydrophilic, released from hCG N-linked glycans which were mixed with an organic matrix (DHB: HCCA) and spotted on the target plate to minimize ion suppression as presented in Figure 3.9. This type of silica-based stationary phase is called Hypercarb® by ThermoElectron Corporation since 1988. To make Hypercarb tips used for glycans purification in this study, the silica was impregnated with a homogenous combination of phenol and hexamine and then heated to 80–160°C to speed up polymerisation. The lengthy process involving immobilisation of the carbon structure, by thermal treatment (2340°C) to achieve graphitisation using argon and cooling to 1000°C, by replacing argon with hydrogen deactivated the surface and made it more homogeneous, hence the end outcome was porous graphitic carbon (Bapiro, *et al.,* 2016). As a result, chromatographic behaviour of PREG (polar retention effect on graphite) was prepared to increase retention and therefore separation of highly polar compounds, such as carbohydrates, and compounds with several hydroxyl, carboxyl and amino groups, which are difficult to retain on conventional alkyl-silica phases (West, *et al.,* 2010).

Next, MALDI-TOF electrospray soft ionisation technique with low-energy collision-induced dissociation (CID) was used to determine a sequence of proteolytic fragments of glycoprotein (neutral glycans), which yield intense signals in the positive reflector mode with sodium, potassium or lithium adducts [M+Na$^+$, M+K$^+$, M+Li$^+$]. Multiple variations in the abundance of antennary fucosylated and sialylated hCG N-glycome have been detected. Amount of sugars (varying degree of site occupancy), sugar composition, glycan sequence, monosaccharide branching i.e. mon-, bi-, tri-, and tetra- antenna and interglycosidic connections were all analysed as it was required for thorough structural elucidation of glycans (Morelle and Michalski, 2007). On average in three hCG standards: AC048 (Fitzgerald), C0434 Batch 1 and C0434 Batch 2 (Sigma) we detected five monoantennary glycans, fifteen biantennary, seventeen triantennary and ten tetrantennary using MALDI-TOF (Table 3.12). Thirty-one glycans were core fucosylated, with fifteen antenna fucosylation out of which seven had multiple antenna fucosylation. We also detected ten singly sialylated and eleven multiple sialylated glycans (Tables 3.7 and 3.8). All these glycans moieties were composed of a relatively small number of different monosaccharides (from Hex3HexNAc2 to Hex7HexNAc6NeuAc4dHex1) which gave rise to multiple peaks with a mass difference of 162 Da for Hexose, 203 Da for HexNAc, 291 Da for sialic acid and 146 Da for deoxyhexose (fucose) on mass spectra. In addition, sialylated glycans found on hCG were more challenging to study using MALDI-TOF-MS as they might produce a combination of ions such as [M-nH + (n + 1)Na]+ and [M-nH + (n + 1)K]+ or [M-nH + (n + 1)Li]+ (Figure 3.7). Thus the m/z of detected N-glycans was established at between 910.1Da to 3758.1Da (Table 3.8). We also observed that by using MALD-TOF, the underivatized sialylated glycans were more prompt to lose a considerable amount of sialic acid during ion extraction or in the ion source, resulting in potential glycan profile distortion (Liu, *et al.*, 2014). Thus to calculate the mass of underivatised glycan motifs with terminal groups: the non-reducing end (1.0078 Da) or reducing end (17.0027 Da) the monoisotopic mass value for sodium (Na$^+$) is 23 Da or potassium (K$^+$) is 3 9Da needs to be added to the sum of the residue masses, minus hydrogen (Morelle and Michalski, 2007). We also compared our results to those of the LC-MS Q Exactive Orbitrap mass spectrometer. This instrument confirmed the structures of sixteen glycan structures detected by MALDI-TOF between 1860 Da and 2319 Da. However, it was not able to detect the fifty glycans detected by MALDI-TOF. As a consequence of the results, the MALDI-TOF technique performed better in detecting more glycan structures, and it was then applied to clinical samples.

Additionally, as described in this work detection of N-linked glycans from hCG using mass spectrometry supports the finding of previous authors including Nemansky *et al.*, (1995) and Gervais *et al.*, (2003) indicating that hCG purified from urine samples contain multiple site-specific N-linked glycans, which are branched (antennas) and at various levels sialylated and/or fucosylated. Further comprehensive structural heterogeneity investigations of hCG glycoproteins started in the early 2000s by Valmu *et al.* (2006), Toll *et al.* (2006), de Medeiros and Norman (2009), Cole (2010, 2012), Fournier (2016) among others and continue till the present day, they detected variation of mon-, bi-, tri-, and tetra- antennary glycan forms on hCG with numerous sialylations and fucosylation moieties using MALDI-TOF and high-performance liquid chromatography (HPLC). Mentioned studies confirmed previous findings and validated the discovery of hCG N-glycans which varied in size from 2149.0 Da to 5859.4 Da (Valmu *et al.* 2006), 1421.0 Da to 2207.0 Da (Toll, *et al.,* 2006), 1257.4-3620 Da (Jansen, *et al.,* 2016) or recently 1579-3923 Da (Ibeto *et al*, 2020) depending on derivatisation and ionisation method. Importantly, the composition of all glycans motifs was comparable in all publications and thesis presented. Also, the degree of structural information desired determines which MS analysis approach is employed for the structural characterization of protein glycosylation.

In summary, we have demonstrated that with a developed methodology glycans can be detached, purified and detected via MALDI-TOF which allowed the characterisation of a wide repertoire of N-linked glycans from human chorionic gonadotropin. N-linked glycan repertoire varied between used hCG standards which were all purified from the pooled pregnancy urine (but contained multiple gestation weeks). This finding has enormous potential for the identification of novel biomarkers in biological samples from healthy pregnancy and pathological urine samples. Also, it is based on the definitive presence or absence of the hCG molecule (confirmed by m/z ratio), which makes it simple and effective.

4. Glycosylation profiling in normal pregnancy and choriocarcinoma

In previous chapters, it was explained how hCG plays a significant role during healthy pregnancy as well as in pregnancy disorders (Chapter 1), also developed and optimised methods for isolating hCG glycoforms from complex biological fluids (Chapter 2) and characterised N-linked hCG glycans (Chapter 3) were presented. It was also emphasized how protein glycosylation modifies physiological and pathological functions in healthy cells, and how any defects or abnormalities in glycosylation may impact disease development and progression (Reily *et al.,* 2019). Consequently, methodologies verified in a study of hCG purification and N-linked glycans characterization by MALDI-TOF mass spectrometry were applied for clinical urine samples from healthy and choriocarcinoma patients.

Human chorionic gonadotropin is a heterodimeric glycoprotein hormone with considerable molecular heterogeneity that required comprehensive examination and evaluation in order to understand its structural characteristics, biosynthesis, functions and pathophysiology (Nwabuobi, *et al*., 2017). During a healthy pregnancy, significant hCG synthesis and secretion start in the first few weeks of gestation. Before the influx of maternal blood into the intervillous space, mononucleated villous cytotrophoblasts (VCT) that cover the floating chorionic villi fuse to form multinucleated syncytiotrophoblasts (ST), which secrete a large amount of hCG. Concomitant with this is a significant change in the trophoblast environment, including an increase in oxygen levels from roughly 2.5 % to 8% after 12 weeks of amenorrhea (Cocquebert *et al*., 2012). Cocquebert, *et al*., (2012) clarified that this oxygen increase may be due to the increased flow of maternal blood to the intervillous space and might explain the maternal peak of hCG in urine and plasma during the first trimester of pregnancy. Additionally, secreted hCG binds to LH/hCG receptor on the surface of mononucleated villous cytotrophoblasts, activates adenylate cyclase, phospholipase C and ion channels, which control second messengers such cAMP, inositol phosphates, Ca^{2+}, and others, which stimulates trophoblastic cell fusion (Frendo *et al*., 2003). After implantation hCG stimulates progesterone secretion by the corpus luteum, maintains progesterone and leptin production, promotes angiogenesis in the uterine endothelium and promotes the formation of the syncytiotrophoblasts (Maymó *et al.,* 2012, Fournier, 2016). Several hCG variants with different structural characteristics are detected in pregnancy urine such as intact hCG (hCG), free β-subunit (hCGβ), free α-subunit (hCGα) but also nicked free β-subunit (hCGβn) or the β-core fragment (hCGβcf) among others (Cole, 2010). The α-subunit is encoded by a single gene on chromosome 6q21.1-q23. The genes located on chromosome 19q13.32, the β subunit molecule is encoded by any of the six

nonallelic genes *CGB8, CGB7, CGB5, CGB3, CGB2,* and *CGB1. CGB1* and *CGB2* are two pseudogenes that might encode a protein unrelated to hCG, whereas the rest of the *CGB* genes are divided into two subtypes: type 1 (*CGB 7*) and type 2 (*CGB 3, -5, and -8*), that vary by three amino acids (Cocquebert, *et al.*, 2012). The elevated levels of total hCG including intact hCG and hCGβ were detected in rare kinds of malignant tumours including choriocarcinoma which is composed of mononuclear cytotrophoblasts and multinucleated syncytiotrophoblasts (Nagai, *et al.,* 2018). Choriocarcinoma can develop after pregnancy, as part of germ cell tumours, or in conjunction with a poorly differentiated somatic carcinoma, each with its own set of clinical characteristics (Cheung *et al.,* 2009).

Furthermore, key cellular processes controlling various physiological and pathological processes in pregnancy and gestational trophoblastic diseases are related to the changes in post-translation modifications such as glycosylation. N-linked glycosylation occurs during target protein translation by the attachment of glycan structures to the amino group of asparagine in ER and Golgi apparatus. Glycans may then be subjected to a variety of extensive modifications, such as sialylation, fucosylation, galactose and GlcNAc addition, and more, resulting in very complex and heterogeneous structures (Oliveira-Ferrer *et al.,* 2017).

Therefore, the heterogeneity of hCG glycoforms detected in pregnancy and choriocarcinoma range from hypo- to hyperglycosylated hCG structures (Cole, 2009). Interestingly, described above alterations in hCG glycosylation are associated with different biological effects e.g. sialylation of gonadotropic proteins such as hCG showed decreased bioactivity and changes in its half-life (Singh, *et al.,* 2018) or fucosylation has been demonstrated to have a significant impact on cell growth and differentiation (Wang *et al.,* 2006). Also, cell adhesion, migration and signalling, as well as immune recognition and host-pathogen iterations can be regulated by glycans (Pinho and Reis, 2015). Conversely, observed changes in increased fucosylation and increased number of antennas (branching) are a signature of malignant cell transformation which are regulated by a family of GlcNAc transferases and contribute to many malignant events (Muthu *et al.*, 2020).

The results presented in this chapter support and augment the above findings by showing dissimilar glycomic profiles of hCG purified from healthy pregnancy with advancing gestation and choriocarcinoma. The 212 N-linked glycans purified from the 32 urine samples were tested. We identified and characterised 83 individual hCG glycan configurations including mono-, bi-, tri- and tetra- antennary from human urine. Additional structural features identified in urine samples included, fucosylation, sialylation, bisected N-glycans, high mannose, sulphated and N-linked glycans with Lewis X terminal epitopes. As a result, our developed and

optimised glycomic-based mass spectrometry methods to purify, detect and characterise, different glycosylation patterns on hCG purified from healthy pregnancy and choriocarcinoma were applied and N-linked glycans structural moieties were compared between healthy and pathological urine samples.

4.1 Materials and Methods

4.1.1 Equipment and Reagents

MALDI -TOF MS/MS (Autoflex)	Bruker, UK
MTP 384 ground steel ID210 1000476	Bruker, UK
Flex Analysis – GlycoQuest software	Bruker, UK
Pipettes	Brand, Germany
Centrifuge	Sigma Aldrich, UK
Vacuum concentrator	Eppendorf, UK
Heating block	ThermoFisher, UK
Balance	ThermoFisher, UK
Vortex mixer	ThermoFisher, UK
Magnetic Rack	ThermoFisher, UK
Peptide-N-Glycosidase F also known as PNGase F	Roche, UK
Nuclease-free water	Invitrogen, UK
hCG beta DuoSet ELISA, DY9034-05	R&D, UK
Alpha-Cyano-4-hydroxycinnamic acid matrix substance for MALDI-MS, 99.0% (HPLC)	Sigma Aldrich, UK
Glycans standards mix: NGAF2 (Hex3HexNac4Fuc1) FA2G1 (Hex4HexNAc4Fuc1) A2F ((Hex5HexNAc4Fuc1NAc2) A3 (Hex6HexNAc5NAc3)	Ludger, Abingdon, UK

4.1.2 Normal pregnancy urine and choriocarcinoma urine samples

Urine was collected, processed, and stored following the same experimental protocols as in Chapter 2, approved by Ethics Committee at Middlesex University - Natural Sciences Ethics Sub-Committee (NSESC) - Application for Ethical Approval - *NSESC Approval for project no 1766*, and performed in accordance with relevant guidelines and regulations. hCG was isolated from morning urine samples (100-500 mL) collected from seven healthy pregnant women subjects at various gestational ages. In addition, urine was collected from six women diagnosed with choriocarcinoma (1-2 mL each). Choriocarcinoma may arise following normal pregnancy, ectopic pregnancy, hydatidiform mole (50%) or rarely from a genital tumour however, for this study the origin of choriocarcinoma was not known. Material Transfer Agreement for the Supply of Human Tissue Materials (MTA Application No 1766) was made by and between Poznan University of Medical Sciences, ul. Polna 33, 61-312 Poznan, Poland and Middlesex University Higher Education Corporation, The Burroughs, London, NW4 4BT. All samples were collected in sterile containers and centrifuged within 1 hour of collection at 5000 rpm for 10 min. Cell-free urine was stored at -80°C.

Table 4.1 Urine pregnancy samples from seven healthy participants. Participants donated samples at various stages of their pregnancy (4-33 weeks).

Pregnancy urine	Gestational age I trimester (4-12 weeks)	Gestational age II trimester (13-25 weeks)	Gestational age III trimester (26-33 weeks)
Participant 1	**+**	**+**	
Participant 2		**+**	**+**
Participant 3			**+**
Participant 4	**+**	**+**	
Participant 5	**+**		
Participant 6		**+**	**+**
Participant 7			**+**

4.1.3 Purification and levels of hCG in urine

Previously described in Chapter 2, Section 2.2.3.3. and Chapter 3, Section 3.5 methods of enrichment with hCG conjugated magnetic Dynabeads and deglycosylation procedure were applied. 50μL of prewashed conjugated anti-hCG beta (5008 SP-5) coated magnetic beads from Medix Biochemica (MedixMAB™), were added to 1mL of pregnancy urine (from weeks 4-33 gestation) or choriocarcinoma urines. The samples were incubated on a rotating platform for 24h at room temperature (RT). Next, beads were immobilized using a magnetic rack, the supernatants were collected and retained hCG was eluted following three additions of 15μL 0.1% TFA as previously described (Section 2.4.3.1). Samples were dried (30°C, 4h), reconstituted in 10μL of 50mM AMBIC and incubated at 100 °C for 10 minutes before the addition of 10μl DTT of 10mM in AMBIC (56°C, 15 min). Next, 10μL of 20 IAA in AMBIC was added and the mixture was incubated for 15 minutes at room temperature. Then PNGase F (Roche, UK) (1 Unit PNGase F for each 0.5μg hCG) was added and incubated overnight at 37°C. Glycans were separated from parent proteins using PGC tips as previously described (Section 3.5). To equilibrate, the tips were flushed twice with 10 μL methanol, followed by 10 μL of 90 % ACN, 0.1% TFA, four times and 10μL 0.1% TFA, five times and the solvent was discarded between each step. Next, the glycans detached from glycoproteins after PNGase F treatment were slowly loaded onto and aspirated from a tip (10 μl, repeated thirty times). Tips were then washed with 10x 10μl 0.1% TFA to remove salts, concentrate sample and then glycans were eluted following the addition of 20 volumes of 10 μl of 50% ACN, 0.1%TFA to elute the glycans. The final 200μl of eluent was left to dry in a rotary evaporator for 6h at 30°C. This was then reconstituted in 4μl of UHQ water and spotted on a 384-ground steel MALDI target plate. 1μl of glycan sample was spotted on the MALDI plate and left to dry (15min, RT), and then 1μl of the matrix (1:1, 10mg/ml HCCA in 50% ACN, 0.1% TFA: 10mg/ml DHB in 50% ACN, 0.1% TFA) was added and left to dry (15min, RT). After the matrix has dried, the target plate was placed in the vacuum source of the MALDI-TOF instrument. Spectra were acquired in a reflector, positive ion mode and chromatogram data were deconvoluted using Bruker's GlycoQuest software (search algorithm in ProteinScape). The GlycoQuest matched m/z MS or MS/MS data against theoretical spectra calculated from databases entries. This works by importing MS data (glycan pool spectra) from the MALDI TOF instrument and selecting a range of search parameters such as type of glycan (hierarchical classification as defined in GlycomeDB database), composition modifications, derivatization type and adducts as previously described in Chapter 3, Section 3.4.

ELISA was used to determine the levels of total hCG (intact hCG and β subunit) in pregnancy and choriocarcinoma samples. The ELISA assay procedure, DY9034-05, Lot number P247055, R&D Systems, UK, comprised of multiple reagents at a specific working concentration, i.e. wash buffer (0.05% Tween20 in PBS), detection antibody reconstituted in reagent diluent (500ng/mL), standard hCG reconstituted in reagent diluent (7.81-2000 pg/mL), streptavidin-HRP diluted in reagent diluent (1:40) and capture antibody reconstituted in PBS (120µg/mL) for coating a microplate (100µL per well, RT, overnight) all based on the manufacture protocol. After washing a 96-well microtitre plate (Maxisorp, Thermofisher, UK) three times with 300µL wash buffer, reagent diluent (300µL) was added and incubated for 2 hours at room temperature to reduce non-specific binding, next it was washed three times more. 100µL of biological samples (diluted with reagent diluent at 1:1000, 1:10 000 and 1:100 000), hCG standards and controls were added to the individual wells, following 2 hours incubation and wash (x3). Next 100µL of the detection antibody (2 hours, RT), and 100 µL of streptavidin-HRP (20 min, RT) were added, followed by washing a plate twice with wash buffer before each step. Finally, 100 µL of substrate solution (20min, RT) and 50µL of stop solution (1N sulphuric acid) were added and the absorbance was read at 450nm within 15min.

4.1.4 Detection of hCG N-linked glycans

Analysis of hCG N-linked glycosylation was carried out on a total of 26 samples isolated from the urine of seven pregnant women and six subjects with choriocarcinoma. Details of gestational age of collected urine samples and hCG concentration were shown in Table 4.1. The concentration of purified hCG glycoforms from urine samples was measured using hCG beta DuoSet ELISA, DY9034-05, R&D, UK. The N-glycan profiles of all hCG samples derived from pregnancy conditions were determined using a glycomic strategy previously optimized in Chapters 2 and 3. This hCG purification and characterisation method based on MALDI-TOF MS analysis was used to determine the hCG glycomic profile from the total population of deglycosylated (PNGaseF) N-glycans found in pregnancy and choriocarcinoma urine. Glycan composition of glycans was defined by deconvoluting MALDI-TOF MS data using GlycoQuest software and all results were additionally manually analysed with the assistance of a glycobioinformatics tool, GlycoWorkBench (Version 2) (Ceroni, et al., 2008). To characterise alleged hCG N-glycan structures collection found in pregnancy and choriocarcinoma urine, underivatized (label-free) glycans were detected based on the intensities of extracted precursor ion chromatograms.

4.1.4.1 hCG glycans identification by GlycoQuest

The MALDI-TOF mass spectrometer was run in reflectron positive ion mode with a monoisotopic mass setting to detect glycans detached from hCG. The accelerating voltage was 29 kV. The glycans data were acquired using the Autoflex MALDI-TOF mass spectrometer (Bruker, UK) and GlycoQuest (Version: 1.3) with internal search method name: MG Glycan MALDI MS AA CarbBank using Glycome DB database, glycan structures were underivatized with modification set to Sulphate (0-4), N-Sulphate (0-1), ions: H+(0-1), Na+(0-2), K+(0-1), Li+(0-1) and MS tolerance (m/z) to 0.7 Da. In the experiment, the glycan calibration standard (internal mixture of four glycans, NGAF 2, FA2G1, A2F, A3) from Ludger, Abingdon, UK was used.

4.1.5 Analysis of hCG protein using MALDI-TOF

hCG and deglycosylated hCG purified from pregnancy and choriocarcinoma urine samples (Section 4.1.2 and 4.1.3) were analysed. All samples were boiled for 10 min and 20 µL of each sample with the addition of 10mM DTT (10µL, 56°C, 10min), 20mM IAA (10µL, RT, 10min) and trypsin 20:1 ratio (20µg/mL, 37°C, 19h) accordingly, following incubation at 37°C, overnight. The trypsin solution was prepared by adding 20µL of 50mM AMBIC combined with 2% acetonitrile as before (Chapter 2, Section 2.2.2). The supernatant was collected, mixed with alpha-cyano-4- hydroxycinnamic acid matrix (1:1) spotted on the MALDI-TOF target plate and air-dried.

4.1.6 Statistical analysis

Data acquired from N-glycans detected in pregnancy and choriocarcinoma urine samples using MALDI-TOF was analysed using Excel Microsoft and MiniTab software. Also, multiple comparisons of hCG concentrations established using ELISAs were carried out with a one-way analysis of variance (ANOVA). The MALDI technique provides a signal intensity for each charged species detected. However, the intensity of the signal is extremely variable. In complex samples such as those derived from biological samples the main cause of this is the appearance of "hot spots" or "sweet spots". The reasons for sweet spot formation are not completely understood. They may be due to variations in the amount and the crystallization of the matrix which may relate to the local analyte concentration. The orientation of the matrix-analyte crystals relative to the spectrometer axis may also be important here (Szájli *et al.*, 2008). For this reason, we have not used quantitative measurements from MALDI based on signal intensities. Instead, we have analysed variations in the abundance of specific glycan species by

comparing their presence or absence in different samples (qualitative approach). Using a chi^2 test, multiple comparisons of the presence or absence of certain glycans in urine samples were conducted, and any variations in the relative proportions of distinct hCG glycan structures across different trimesters of pregnancy and with choriocarcinoma were detected. In addition, binary logistic regression analysis with backward elimination was used to determine any specific structural variations that were significantly different at different gestational ages or with choriocarcinoma. The odds ratio (with a 95% confidence interval) was calculated to determine how likely a specific glycan species was to appear from choriocarcinoma samples compared to the first trimester of normal pregnancy. Because choriocarcinoma occurs in early pregnancy when cells that were formerly part of the placenta become malignant, this was the most relevant comparison of choriocarcinoma N-linked glycan composition with first-trimester pregnancy. Choriocarcinoma can develop as a result of a miscarriage, abortion, ectopic pregnancy, or molar pregnancy, which occurs when an egg is fertilised but the placenta develops into a mass of cysts rather than a fetus (Holland and Hreshchyshyn, 2012). For all tests, p values were considered significant at a 95% confidence interval (p≤0.05).

4.2 Results

A total of 212 N-linked glycans were detected from the 32 urine samples tested. We identified and characterised 83 individual hCG glycan configurations including mono-, bi-, tri- and tetra-antennary (5, 56, 11 and 11, respectively) from human urine. Additional structural features identified in urine samples included, fucosylation (38 non-fucosylated; 32 core and 13 antenna fucosylation), sialylation (15 single, 17 multiple and 51 with no sialic acid attached), 16 bisected N-glycans, 6 high mannose, 9 sulphated and N-linked glycans with Lewis X terminal epitopes. In the first trimester of pregnancy, a significant difference in the size distribution of N-glycans in comparison to choriocarcinoma was observed. Also, there was an association between hCG concentration and the number of glycans detected (Pearson correlation, r=0.85, p<0.005) and no significant difference in hCG concentration between the first, second and third trimester.

4.2.1 hCG concentration in urine samples

hCG was detected in all urine samples collected from pregnancy urine and choriocarcinoma subjects (Table 4.2). There was a difference in hCG concentration, i.e third trimester was significantly higher than the first trimester (t-test, p=0.03). However, there was no significant difference between the first trimester of pregnancy and choriocarcinoma samples (t-test, p=0.39).

Table 4.2. Average hCG concentration in urine from pregnancy (first, second and third trimester) of seven healthy participants and six choriocarcinoma patients were measured using ELISA, DY9034-05, R&D Systems, UK. CV- coefficient variation

Sample	hCG concentration [µg/mL] (mean ± SEM)	CV [%]	Average glycan number
1st-trimester pregnancy	1270.8 ± 433.5	102	16
2nd-trimester pregnancy	1023.6 ± 581.8	127	20
3rd-trimester pregnancy	2617.1 ± 684.0	52	27
Choriocarcinoma	897.7 ± 106.9	17	19

4.2.2. Glycan composition in normal pregnancy urine

The N-linked glycans were detected in all the urine samples tested. There was a significant variation in the numbers of glycans detected among samples (between 7 and 34 glycans per sample). The average number of glycans increased with gestational age and there was a significant increase in third-trimester urine compared to the number in the first trimester (ANOVA with Tukey post hoc $p < 0.05$, Figure 4.1). On average 24 and 19 N-linked glycans were detected in the pregnancy (first, second and third trimester) and choriocarcinoma samples respectively. The 11 most numerous glycan structures in healthy and choriocarcinoma urine were presented in Table 4.3.
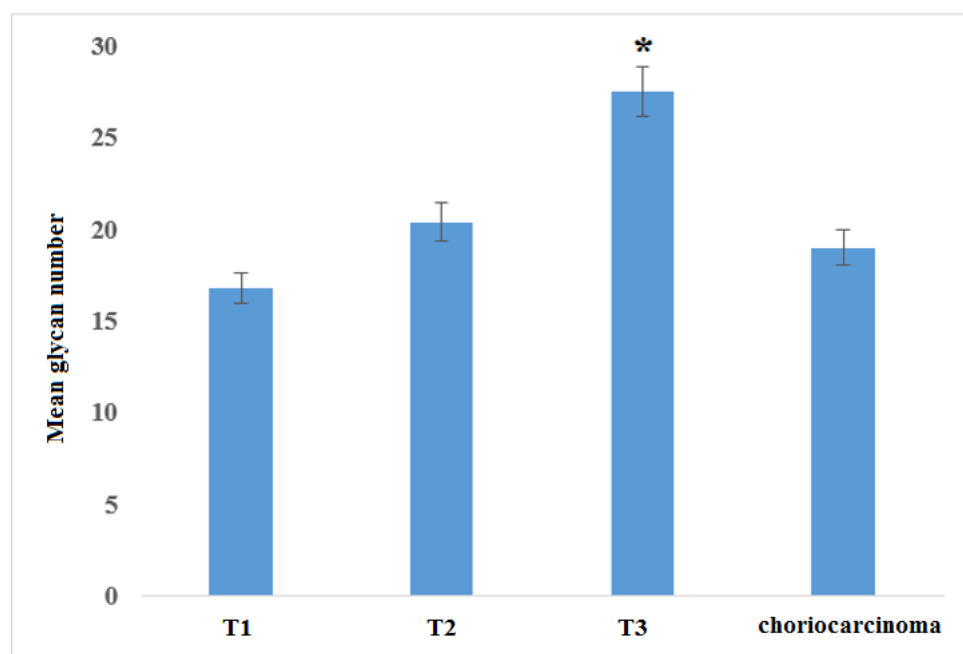
Figure 4.1 The mean number of glycans in each of the urine samples collected from the T1, T2 and T3 trimesters of pregnancy and choriocarcinoma detected using MALDI-TOF MS with GlycoQuest software. Error bars represent mean ± SD. * represent significant difference compared to first-trimester urine (ANOVA with Tukey post hoc test ($p < 0.01$)).

Table 4.3 Most abundant glycan m/z and structure detected in pregnancy samples. Hex-Hexose, HexNAc- N-Acetylhexosamine, dHex–dHexose, NAc–sialic acid (N-acetylneuraminic acid)

Glycans m/z	Structure
934	Hex3HexNAc2
1095	Hex3HexNAc2dHex1
1111	Hex4HexNAc2
1136	Hex3HexNAc3
1225	Hex4HexNAc2dHex1
1355	Hex3HexNAc4
1412	Hex3HexNAc3dHex2
1420	Hex6HexNAc2
1502	Hex4HexNAc4
1519	Hex3HexNAc5
1694	Hex6HexNAc3NAc1
1719	Hex4HexNAc3dHex1NAc1
1922	Hex4HexNAc4dHex1NAc1
2222	Hex3HexNAc6dHex2
2558	Hex6HexNAc7dHex1
2903	Hex7HexNAc5dHex1NAc2
3167	Hex11HexNAc6dHex1

Figure 4.2 showed that there was no difference in the size distribution of glycans detected with advancing gestation.
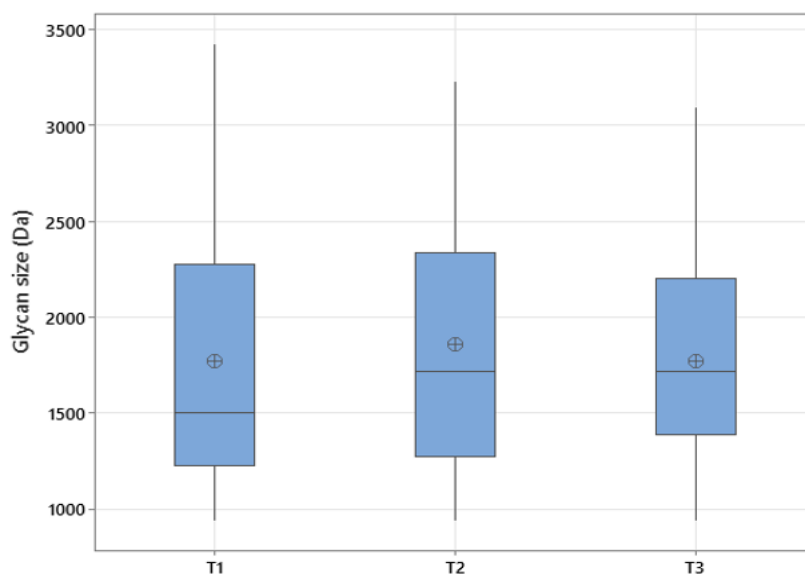
Figure 4.2 Size distribution of glycans detected following extraction of hCG collected from urine from subjects in the first, second and third trimesters of normal pregnancy (T1, T2 and T3).

Figure 4.3 illustrated the distribution of single and multi-antenna glycans detected from hCG in normal pregnancy with advancing gestation. There was a decrease in single antenna glycans and an increase in multi-antenna glycans (>1antenna) with advancing gestation. However, this difference did not reach statistical significance (chi^2 test p=0.319).
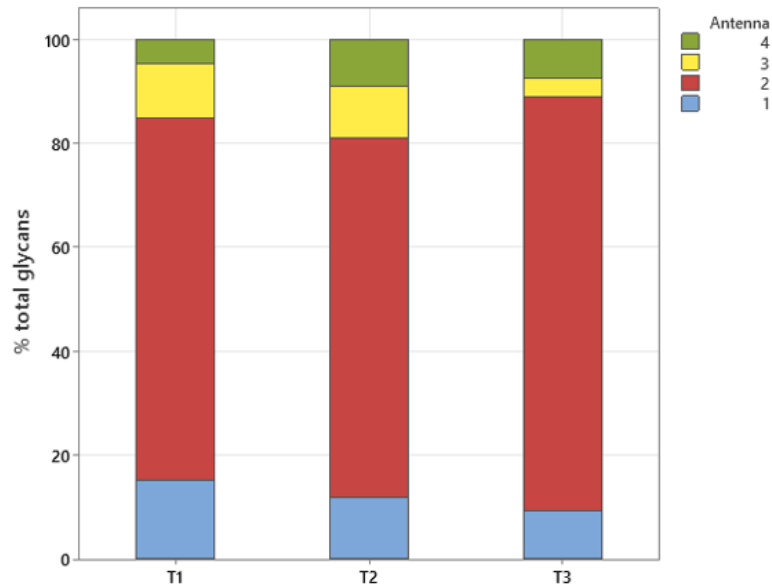
Figure 4.3 Distribution of single (1) and multi-antenna glycans (2, 3, 4) detected from hCG in normal pregnancy with advancing gestational age. The y-Axis values are percentages of the total number of glycans.

As normal pregnancy advanced there was a decrease in non-fucosylated glycans and a concomitant increase in the proportion of multi fucosylated hCG N-linked glycans (Figure 4.4). This difference achieved statistical significance (chi^2 test, p=0.01). The proportion of single fucosylated species remained relatively stable at each trimester. Also, the y-Axis values are percentages of the total number of glycans in subsequent figures.
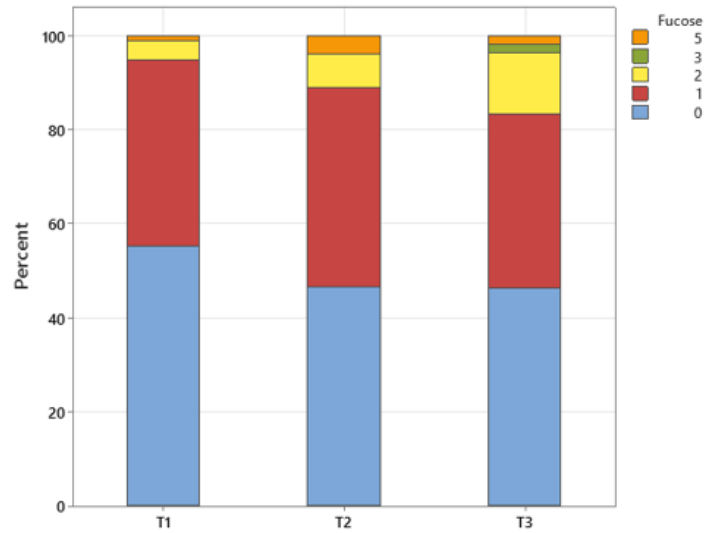
Figure 4.4 Distribution of non-fucosylated (0), single (1) and multi-fucosylated glycans (2, 3, 5) detected from MALDI and GlycoQuest analysis of glycans extracted from hCG from urine of first, second and third trimesters of pregnancy (T1, T2 and T3 respectively). Significant difference in the expression of multifucosylated glycans was observed.

The number of sialic acids incorporated into each glycan showed a decrease as pregnancy advanced and in the third trimester, there were no tri-sialic acid glycans (Figure 4.5). However, this difference did not achieve statistical significance (chi^2 test, p=0.765).
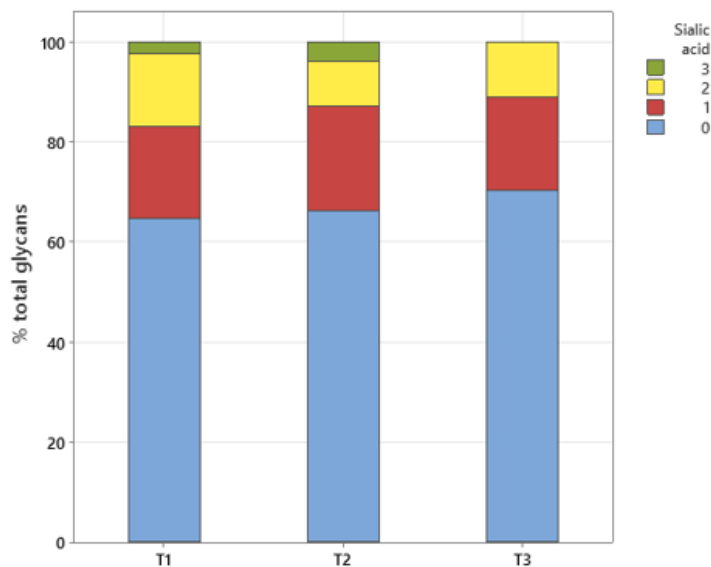
Figure 4.5 Distribution of non- sialylated, single and multi- sialylated glycans detected from MALDI and GlycoQuest analysis of glycans extracted from hCG collected from the urine of subjects in the first, second and third trimesters of pregnancy (T1, T2 and T3 respectively).

Finally, there was an increase in the proportion of bisected glycans with advancing gestation. However, once again, this difference did not achieve statistical significance (chi^2 test p=0.584) (Figure 4.6).
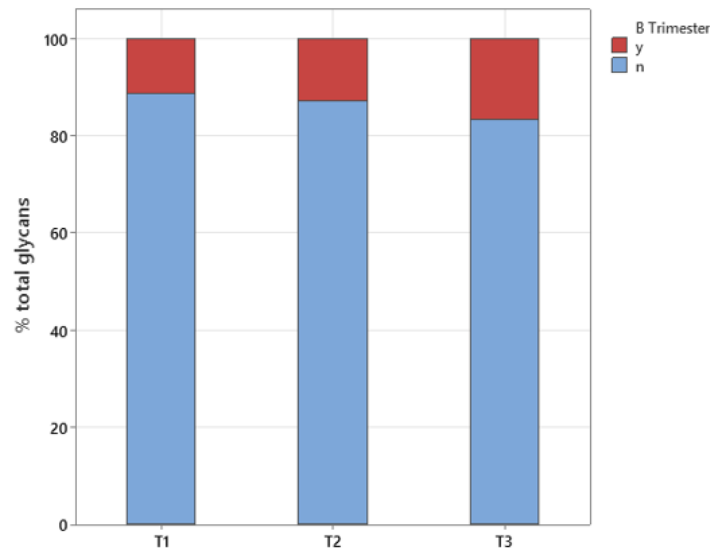
Figure 4.6 Distribution of bisected glycans (red= yes; blue=no) detected from MALDI and GlycoQuest analysis of glycans extracted from hCG collected from the urine of subjects in the first, second and third trimesters of pregnancy (T1, T2 and T3 respectively).

4.2.3 Glycan composition in choriocarcinoma urine

There was a significant difference in the distribution of glycans size in first-trimester urine and choriocarcinoma with an increase in high molecular weight glycans in choriocarcinoma compared to the first trimester of pregnancy (chi^2 test, p=0.005). The size distribution of glycans in the first trimester of normal pregnancy and choriocarcinoma is shown in Figure 4.7.
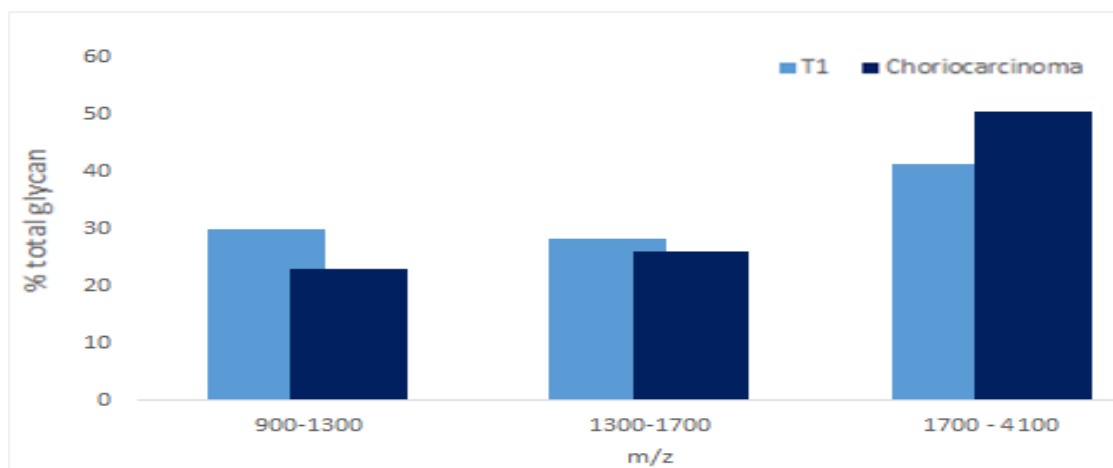
Figure 4.7 Histogram showing the size distribution of free non-reduced N-linked glycans from first-trimester pregnancy and choriocarcinoma urine samples, detected using MALDI-TOF MS. Mass range between m/z ≥900>4100 Da was characterised using MALDI-TOF MS with GlycoQuest software.

In choriocarcinoma, 30 % more of the glycans were tri-and tetrantennary glycans in comparison to a healthy pregnancy (Figure 4.8). Twice as many tri-and tetrantennary glycans were detected in choriocarcinoma (chi^2, p=0.004), conversely to a 10% decrease of mono- and di-antennary.
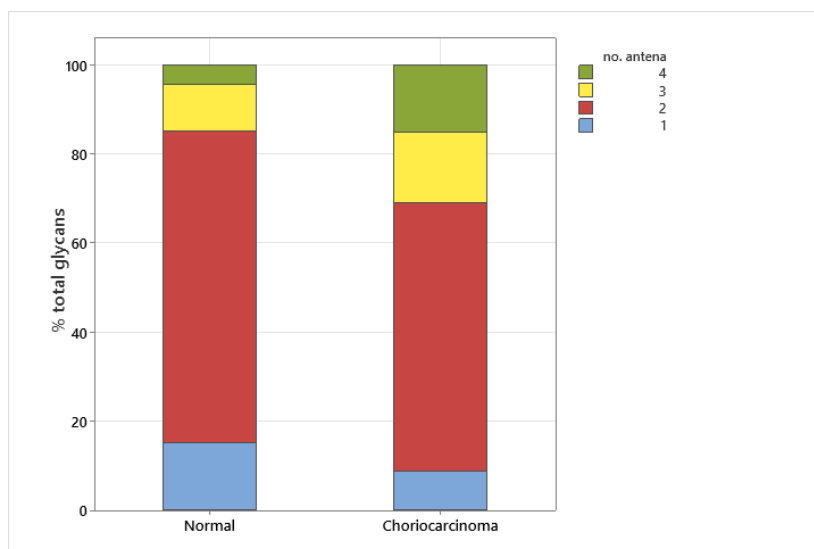
Figure 4.8 Distribution of mono and multi-antenna glycoforms extracted from hCG in urine collected from normal pregnancy (first-trimester) and choriocarcinoma and characterised using MALDI-TOF MS with GlycoQuest software.

The proportion of multi fucosylated glycans (including 2, 3, 5 and 7 fucose groups) was three times higher in choriocarcinoma compared to the first trimester of normal pregnancy (chi^2, p≤0.001) (Figure 4.9).
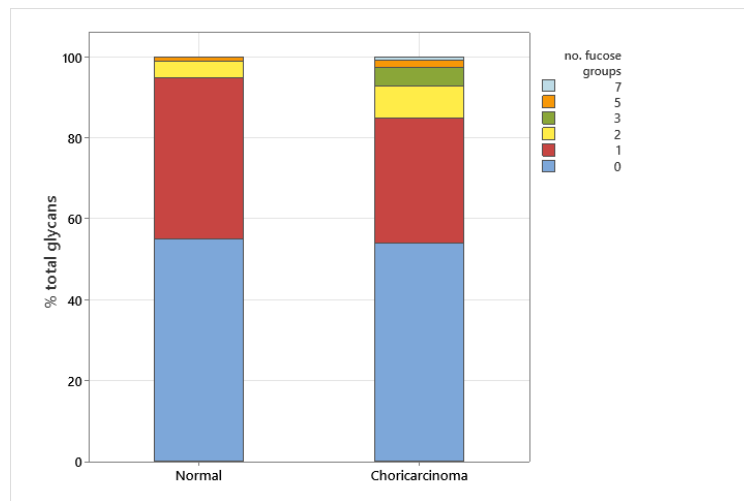
Figure 4.9 Number of fucose residues detected in the first trimester of normal pregnancy (T1) and choriocarcinoma urine characterised using MALDI-TOF MS with GlycoQuest software.

Multiple core- or antennal sialylations (two, three or four sialic acid residues) were also detected in choriocarcinoma urine and four sialic residues were only detected in choriocarcinoma.
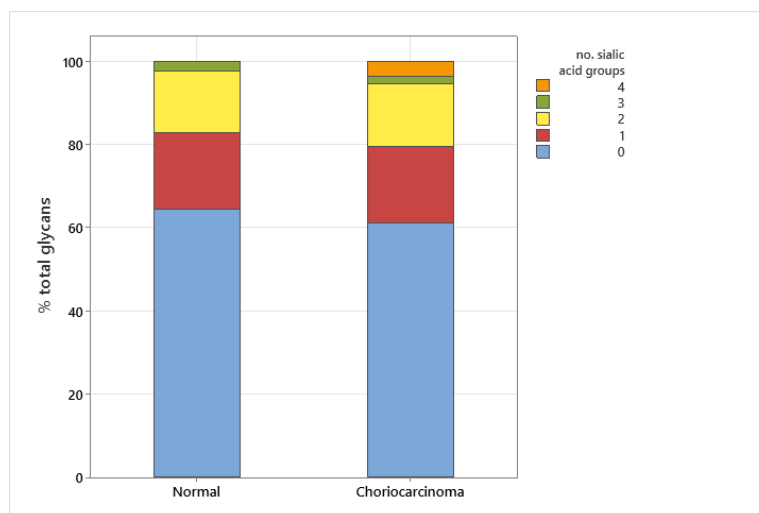
Figure 4.10 Distribution of sialic acid residue number on glycans extracted from hCG in urine samples collected from normal pregnancy (first trimester) and choriocarcinoma. characterised using MALDI-TOF MS with GlycoQuest software.

The bisecting GlcNAc transferred to the core mannose residue of complex or hybrid N-glycans were detected in choriocarcinoma and normal pregnancy urine samples. Although significantly more bisected N-glycans were detected in urine from cancer patients (chi^2 test, p=0.03), as presented in Figure 4.11.
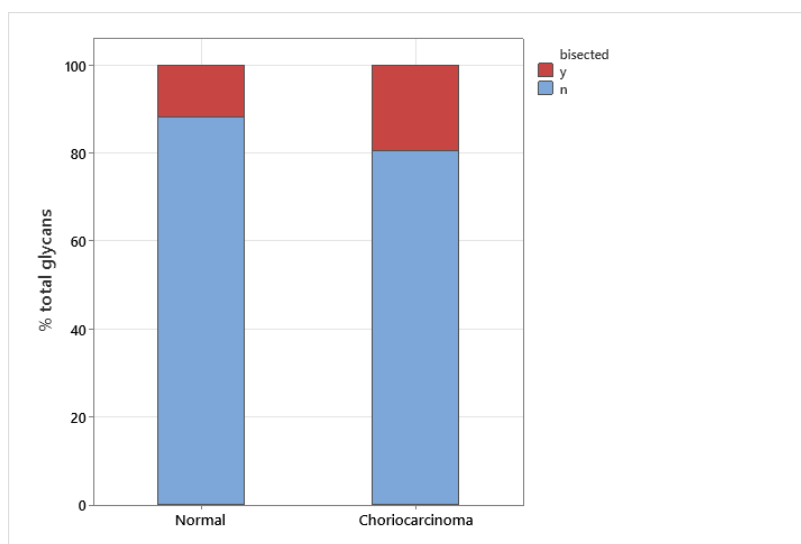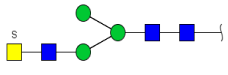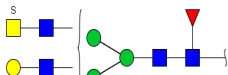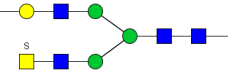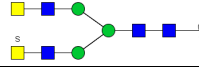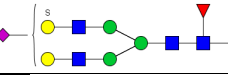
Figure 4.11 Distribution of bisected and non-bisected n-glycans detected in hCG extracted from normal pregnancy (first-trimester) and choriocarcinoma urine using MALDI-TOF MS with GlycoQuest software.

We were also able to detect changes in N-glycan structure accordingly to the number of the sulphate group present. Interestingly, N-glycan sulphation was not seen in first and second-trimester samples but was observed in late pregnancy and also in two out of six choriocarcinoma patients. The five sulphated glycan structures extracted from third-trimester pregnancy urine and choriocarcinoma urine are shown in Table 4.4.

Table 4.4 Sulphated N-glycans detected in late pregnancy (week 33) and choriocarcinoma urine samples using MALDI-TOF MS with GlycoQuest software. As before: green circle=manosse, yellow circle=galactose, blue square=HexNAc, red triangle=core-fucose, purple diamond=sialic acid and yellow square with S letter on top shows single or multiple sulphated groups.
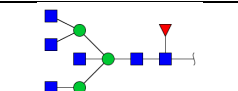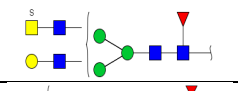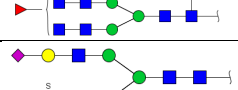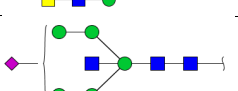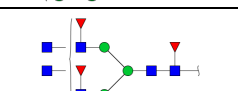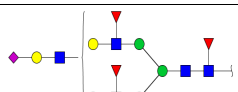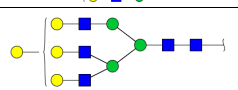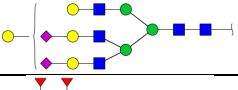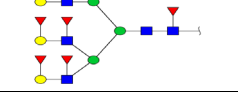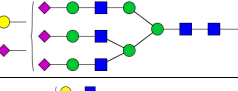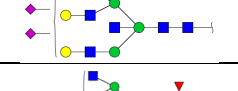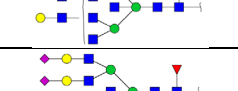
N-linked glycans composition and structure in late pregnancy		N-linked glycans composition and structure in choriocarcinoma	
Hex3HexNAc4		Hex4HexNAc5dHex1	
Hex5HexNAc4dHex1		Hex4HexNAc5NeuAc1	
Hex3HexNAc6			
Hex5HexNAcNeuAc1dHex1			
Hex3HexNAc6dHex1			

4.2.3.1 N-glycans unique to choriocarcinoma urine

We were also able to identify fourteen N-glycans in a pooled urine sample from six choriocarcinoma patients that were not detected in normal pregnancy urine. As mentioned before in Section 3.5 (page 117), each clinical sample was only analysed once due to the limited amount of experimental material. The combined glycans structures from all of these samples and representation of individual structures from choriocarcinoma patients 1-6 were presented in Table 4.5. Once again, due to the MALDI-TOF limitations, it was very important to mention that all presented structures are postulations only and complete compositional assignments require further experiments, including MS/MS analysis to distinguish between different isomeric structures or MS/MS CID (collision-induced dissociation) and HCD (higher-energy collisional dissociation) to help with linkage position assignment (site-specific N-glycosylation). Also, enzymatic treatment to further characterise glycosyltransferase specificity will be needed for more precise compositional assignments of N-glycans purified from choriocarcinoma urine. Significantly, several of these glycans included characteristics that may be specific to choriocarcinoma glycosylation. Additionally, (multi- fucosylated and

sialylated) Lewis X terminal epitopes were presented in choriocarcinoma samples as shown in Figure 4.12. Additional binary logistic regression tests predicted the relationship between first-trimester pregnancy and choriocarcinoma samples (Section 4.2.4).

Table 4.5 Unique N-glycans characteristic to choriocarcinoma glycosylation which includes fucosylation, sialylation, bisected, sulphated and multi-antennal N-glycan structures. These fourteen structures were observed once or more and combined from six choriocarcinoma urine patients. The possible compositional assignments in the table below are postulations only and have to be treated with caution.

Number	Composition	Structure	Patients					
			1	2	3	4	5	6
1	Hex3HexNAc6dHex1				+	+		+
2	Hex4HexNAc5dHex1							+
3	Hex3HexNAc6dHex2				+	+	+	+
4	Hex4HexNAc5NeuAc1							+
5	Hex5HexNAc3NeuAc1		+				+	
6	Hex3HexNAc6dHex3				+			
7	Hex6HexNAc5NeuAc1dHex3				+			
8	Hex7HexNAc5				+			
9	Hex7HexNAc5NAc2		+					
10	Hex6HexNAc6dHex7		+					
11	Hex7HexNAc5NeuAc4		+		+			+
12	Hex6HexNAc6NeuAc2		+					
13	Hex4HexNAc9dHex1		+					
14	Hex7HexAc7NeuAc4dHex1		+					

Composition	Structure
Hex1HexNAc2dHex2	
Hex2HexNAc2dHex2	
Hex2HexNAc2NeuAc1dHex1	
Hex2HexNAc2NeuAc2dHex2	

Figure 4.12 Chromatogram of sialylated and fucosylated Lewis [X] terminal epitopes detected only in choriocarcinoma urine and characterised using MALDI-TOF MS with GlycoQuest software.

4.2.4 Binary logistic regression

We have demonstrated that choriocarcinoma is associated with an increased proportion of multi-fucosylated, bisected and multi-sialylated glycans as determined from chi^2 testing. Binary logistic regression was carried out to further investigate this phenomenon and determine if these differences will significantly affect the odds ratio. This is the ratio of the odds of the presence of these glycans being present if the subject has choriocarcinoma relative to the odds of the glycans being present if the subject does not have choriocarcinoma (normal pregnancy). The results are presented in Table 4.5.

Table 4.6 Binary logistic regression analysis of glycans characteristics present in urine samples collected from individuals with choriocarcinoma and normal first-trimester pregnancy. OR denotes odds ratio. Glycan features include fucosylation, sialylation, bisected and multi-antennal N-glycan structures.

Glycan Feature	Estimated OR (Choriocarcinoma)	95% confidence interval	P-value
>1 fucose	3.56	1.59 - 8.01	0.002
>2 antenna	2.67	1.52 – 4.69	0.001
bisected	1.80	0.96 – 3.36	0.066
>1 fucose AND >2 antenna	6.80	2.42 – 19.09	0.01
size	1.0001	0.99 – 1.06	0.607

Logistic regression analysis revealed that from this small sample set the most promising biomarker characteristic for the identification of patients with choriocarcinoma compared to normal pregnancy was the presence of more than single fucose attached to N-glycans (OR 3.56). The presence of more than 2 antennas on the N-glycan is also characteristic of choriocarcinoma. Interestingly, in both cases, the 95% confidence was >1 indicating that these 2 structural features may be useful in identifying individuals with choriocarcinoma. Combining these two characteristics further increased the odds ratio (2.42 - 19.09). However, the very low number of glycans fulfilling both criteria (n = 22) means that this finding should be treated with caution, also bisected glycans and glycan size were not predictive.

4.3 Discussion

In this chapter, hCG glycoforms purified from pregnancy and choriocarcinoma urine samples were further characterised using MALDI-TOF and GlycoQuest software. Significant differences between N-linked glycan profiles between the healthy pregnancy and choriocarcinoma urine were distinguished (Table 4.1).

The N-linked glycans were detected in all the urine samples tested. On average (Figure 4.1) 24 and 19 N-linked glycans were detected in the pregnancy (first, second and third trimester) and choriocarcinoma samples respectively. In a healthy pregnancy, there was no difference in the size distribution of glycans detected with advancing gestation (Figure 4.2) however a decrease in single antenna glycans and an increase in multi-antenna glycans (>1antenna) with advancing gestation was observed (Figure 4.3). Also, a decrease in non-fucosylated glycans (Figure 4.4) and sialic acid number (Figure 4.5) as pregnancy advanced, a concomitant increase in the proportion of multi fucosylated hCG N-linked glycans and an increase in the proportion of bisected glycans with advancing gestation (Figure 4.6). Next, 14 unique choriocarcinoma urine N-glycans structures (Table 4.5) were detected and binary logistic regression tests predicted the relationship between pregnancy and choriocarcinoma samples. The results of glycan composition included an increase in molecular weight (>2100Da) glycans which were significantly higher than glycans detected from the first trimester of pregnancy (Figure 4.7). Also, glycans were significantly larger with a 30% increase in tri-and tetra-antennary structures (Figure 4.8). Furthermore, the proportion of multi fucosylated glycans (Figure 4.9) was three times higher and multiple core- or antennal sialylations (two, three or four sialic acid residues) (Figure 4.10) were also detected in choriocarcinoma urine. Additionally bisected, and sulphated glycans (Figure 4.11 and Table 4.4) were detected in urine from choriocarcinoma. As a result, we demonstrated that choriocarcinoma is associated with an increased number of multi-fucosylated, bisected, and multi-sialylated glycans. Furthermore, statistical analysis revealed (Table 4.6) that the presence of more than one fucose attached to N-glycans was the most promising marker for identifying individuals with choriocarcinoma compared to normal pregnancy from this limited sample set. Choriocarcinoma was further distinguished by the presence of more than two antennas on the N-glycan. Interestingly, in both situations, the 95 percent confidence level was more than one, indicating that using both traits might be effective in identifying people with choriocarcinoma in urine analysis. Combining these two structural features boosted the odds ratio even further (2.42 - 19.09). However, due to the small number

of glycans that met both requirements (n = 22), this conclusion should be interpreted with care. Additionally, glycan size and bisected glycans were not predictive.

Interestingly, described above heterogeneity of N-glycans glycosylation purified from glycoproteins was previously detected in several other types of cancers, e.g. in breast cancer, hybrid glycan structures i.e. α-1,3fucosylated, mono-galactosylated triantennary glycans variations included high-mannose, glycans sialylation, fucosylation and branching were detected (Hamid *et al.,* 2008, Scott and Drake, 2019). In ovarian cancer complex, N-glycans were rarely found however high-mannose, bisected, including core fucosylated, galactosylated biantennary glycans and sialyl Lewis[X] were observed (Everest-Dass *et al*., 2016, Trinchera *et al.,* 2017), in prostate cancer high-mannose, sialylation, fucosylation was detected on prostate-specific antigen (PSA) (Butler and Huang, 2021), in colorectal cancer core-fucosylation and α1,3-fucosylated mono-galactosylated triantennary glycans were seen (Doherty *et al.,* 2018) and in gestational trophoblastic diseases (GTD) presence of bisected type N-glycans with abundant Lewis[X] antigens, tri- and tetra-antennary N-glycans and sialylation were described by Ibeto *et al.* (2020).

As a result, presented in this chapter altered glycosylation N-glycans pattern purified from choriocarcinoma urine might be considered a hallmark of malignant transformation. In detail, our findings on glycan composition and structure were also matched with previous glycomic studies indicating that protein glycosylation patterns were associated with pathological adaptations (Hoja-Łukowicz *et al.,* 2018). These include for example increased branching of N-glycans by triggering the transcription of *MGAT5* genes and therefore the expression of glycosyltransferases such as N-acetylglucosamine transferase V (GnTV) (Stowell, *et al.,* 2015). Moreover, overexpression of group of N-acetylglucosamine transferases such as GnT-I, GnT-II, GnT-III, GnT-IV, GnT-V (Figure 4.13) causes abnormal assembly of glycans and compromising cell to cell adhesion and signalling pathways which contribute to tumour invasiveness and metastases (Taniguchi and Korekane, 2011). Furthermore, the formation of tri- and multi-antennary branching structures attached to the core pentasaccharide in the Golgi apparatus was confirmed to be associated with hepatocellular (Fan *et al.,* 2012) and gastric cancer are regulated by overexpressed *MGAT4B* and *MGAT1* genes (Carvalho *et al.,* 2016).
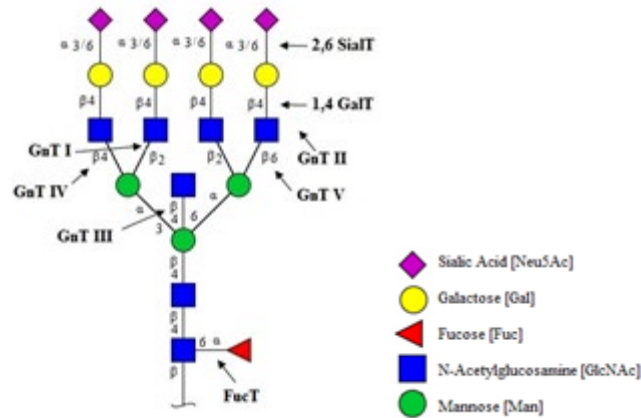
Figure 4.13 N-linked glycan structure with a representation of the key enzymes responsible for the addition of specific sugar residues involved in the N-glycosylation process. Enzymes include N-acetylglucosamine transferases (GnT-I, GnT-II, GnT-III, GnT-IV, GnT-V), N-acetylgalactosamine transferase (β1,4 GalT), sialytransferases (α2-6-SialT) and fucosyltransferases (FucT 8 – core fucosylation) (Chen *et al.*, 2020).

Additinally, Nishino *et al*., (2017) described the close association of N-acetylglucosaminyl-transferase IVa and the formation of β1-4GlcNAc branches on mannose core N-glycans which promotes the invasion of choriocarcinoma and confirms our finding of multi-antennary glycans detection on hCG purified from choriocarcinoma urine (Table 4.5). Moreover, detected bisected structures of N-glycans (Section 4.2) have been previously connected to *MGAT3* gene stimulation of N-acetylglucosamine transferase (GnT III) causing the addition of 1,4 GlcNAc to the beta-linked mannose of the trimannosyl core (three mannoses) (Oliveira-Ferrer, *et al.,* 2017). Next, the transfer of fucose from GDP-fucose to core or terminal (antennal) of N-glycans was described by Zhao *et al*., (2014) as overexpression of *FUT1-FUT11* encoding fucosyltransferases (FucT1 - FucT11). The implications of described variations have been mostly associated with malignant transformations in gastric cancer (Zhao, *et al.,* 2014) and colorectal cancers (Gomes, *et al*., 2013). These include decreased core- fucosylation and increased terminal fucosylation on N-linked glycans coded by *FUT 8* with an increase of sialylated type-2 N-acetyllactosamine (type 2 glycan units) or LacNAc sequences which can be further capped with fucose (α1,2 or α1,3 fucosylation) and sialic acid as shown in Figure 4.14 (Gomes *et al*., 2013, Zhao *et al.,* 2014, Pinho and Reis, 2015). Furthermore, glycans with sulphated terminal β-linked GalNAc were also detected in the urine of choriocarcinoma and pregnancy urine. Previous studies confirmed the detection of sulphated groups in pituitary

glycoproteins such as luteinizing hormone (LH), thyroid-stimulating hormone (TSH), where an alteration in glycosylation regulated the ovulatory cycle and not in follicle-stimulating hormone (FSH), although it is made in the same cells (Cahoreau *et al.,* 2015). A free α subunit common in LH, TSH and FSH was present in all pituitary cells however, the differences in enzyme expression between LH and TSH (*CHST8* or *CHST11*) activate different sulfotransferases and cause variations in the synthesis of sulfated terminal β1-4GalNAc following the addition of α2-6 sialic acid residues. The GlcNAc residues extended with a β1–4GalNAc may also be modified by α1–3 fucose residues however, FSH carries more β1–4Gal residues which are stimulated by β1–4GalT enzymes (Zheng *et al.,* 2008).
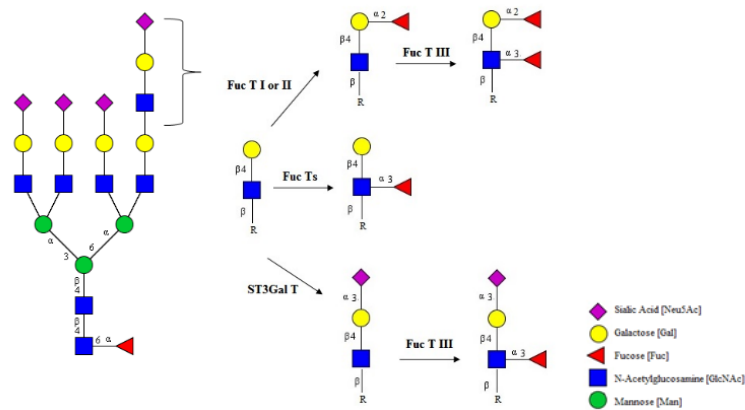
Figure 4.14 Typical glycan extensions including complex N-glycans with poly-N-acetyllactosamines (LacNAc) Type II [Galβ1–4GlcNAcβ1–3]n (of variable length) unit on any branch and Lewis antigens. Fuc T I, II, III- fucose transferase I, II, III; ST3GalT – α2-3-sialytransferase (Adapted from Pinho and Reis, 2015).

This suggests that the characterisation of hCG glycosylation has the potential to distinguish choriocarcinoma from non-pathological conditions. Current diagnostic procedures include ultrasound pelvic examination and hCG concentration checks in urine or blood, which rely on antibody sensitivity and might be ineffective (Duffy *et al*., 2015). As a result, a new testing strategy that includes glycan mass profiling (GMP) from urine samples, which is a fast analysis of glycan composition based on molecular mass using MALDI-TOF mass spectrometry, could provide a more rapid and non-invasive diagnosis. For instance, the changes in expression of glycosyltransferases have already been linked to cancer, particularly as sialylated glycans play

a crucial role in cellular recognition, cell adhesion, cell signalling and cell detachment from the tumour and metastatic cancer spread into the extracellular matrix (ECM) and other cells in the same tumour mass which clinically is associated with poor prognosis of cancer patients (Dobie and Skropeta, 2021). In consequence, it is possible that the identification of an hCG glycosylation profile may also be useful in other disorders which are associated with altered hCG expression. As an example, the alteration in hCG content occurs in both blood and urine and it was previously suggested by Kovalevskaya, *et al*., (1999) that the determination of the relative balance of hCG isoforms has diagnostic application in predicting pregnancy outcomes such as miscarriage, ectopic pregnancy (Homan *et al.,* 2000), pregnancy-related disorders and gestational trophoblastic disease (Berger and Lapthorn, 2016), ovarian and cervical malignancies, colon (Lundin *et al.,* 2001), bladder (Mora *et al.,* 1996), and lung cancer (Yokotani *et al.,* 1997), Down syndrome pregnancy (Cole *et al.,* 1999) and other pregnancy complications (Kovalevskaya *et al.,* 2002). Current diagnostic techniques for these listed above diseases include invasive physical examinations, biopsies, ultrasounds, magnetic resonance imaging (MRI), morphology blood tests and genetic screening test, thus the development of new, early detection approaches for the potential diagnostic method used in the identification of various diseases would be highly beneficial.

Chapter 5

In this thesis, the principal aim was to develop and optimise mass spectrometry-based methods for the identification and characterisation of carbohydrate moieties on human chorionic gonadotropin in biological samples. This methodology has then been applied to the investigation of the structural features of N-linked glycans isolated from hCG purified from urine samples from healthy pregnancies and patients with choriocarcinoma.

5.1 hCG purification and deglycosylation

In Chapter 1 hCG was described as a heterodimeric glycoprotein hormone with considerable molecular heterogeneity (Berger *et al.,* 2002). In order to represent this diversity for the current study, a wide range of commercially available intact hCG standards purified from a pool of pregnancy urine along with well-characterised glycoproteins such as ovalbumin were used. A critical stage in the hCG purification procedure was ensuring that several hCG isoforms were recovered from the urine samples. This was ensured by using a pool of two antibodies known to bind multiple hCG isoforms. A significant effort was focused on the precise and detailed enrichment and characterisation of the hCG molecule prior to the generation of a pregnancy hCG N-linked glycan collection (database). Thus, in Chapter 2 the development of methodologies for the purification/enrichment of hCG isoforms from human pregnancy urine and conditioned media of choriocarcinoma cancer cells are described. All hCG glycoforms from various samples were examined using sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE), hCG isoforms concentration was measured using enzyme-linked immunosorbent assays (ELISA) and final glycopeptide mass fingerprinting (PMF) identification technique using Matrix-Assisted Laser Desorption /Ionisation (MALDI) Time of Flight (TOF) Mass Spectrometry (MS) was applied. As a result, the above methods confirmed the presence of hCG in all tested urine samples by matching its molecular weights (Cole, 2012) or by comparing the amino-acid sequence with the published hCG sequence found in the protein database from NCBI (Elliot *et al.,* 1997).

We focused on the development of hCG enrichment methods from urine samples to increase hCG yield and enhance the detection of hCG using MALDI-TOF MS. Preliminary methods included lectin affinity chromatography (Abbott and Pierce, 2010) and hCG antibody-coupled protein G sepharose beads (Kavran and Leahy, 2014). However, it was found that these methods lacked the ability to capture a sufficient amount of hCG from pregnancy urine. Nevertheless, it was found that hCG antibody-coupled tosylactivated, superparamagnetic

Dynabeads were much more effective and were able to capture 87% of hCG purified from clinical urine samples after 12 hours of incubation at room temperature. To support our findings we compared our results with Levernæs *et al*, (2017) where a similar recovery procedure applied to the hCG molecule achieved around 80% recovery.

Following hCG purification and enrichment from biological samples, glycans were detached from the parent glycoprotein using a four-stage deglycosylation procedure that included denaturation, reduction, alkylation, and enzymatic removal of the glycans described in Chapter 3. During the development of these procedures several factors, including the concentrations of dithiothreitol, iodoacetamide, and PNGase F were optimised. Cleaved hCG glycans were extracted from the glycan/protein mixture using porous graphitic carbon (PGC) Hypercarb tips. This approach was utilised to retain glycans (polar analytes) by charge induction on the polarizable PGC and has proved to be particularly successful for N-glycans analysis, including sialylated, fucosylated, and sulphated hybrid structures (Bapiro *et al.,* 2016, She *et al.,* 2020).

In addition, the sample/matrix application to the MALDI plate was thoroughly examined (Section 3.4.2). This included adjusting different ways of applying the sample and matrix mix onto a MALDI-TOF target plate. Thus, by optimising sample:matrix application in the sample preparation step for MALDI-TOF analysis, a 4 fold increase in the peak intensity (80%) was achieved, allowing the data to be easily reproducible. This indicated that the best conditions were obtained by air-drying sample before applying the layer of matrix. Previous studies in 2007 by McComb and coworkers investigated various sample application procedures, confirming our conclusion that layering a sample onto a target MALDI plate using a specific approach (multi-step sample/matrix) can enhance sensitivity, improve the signal-to-noise ratio, and minimise background noise (McComb *et al.,* 2007).

Optimised conditions permitted successful detection and characterisation of N-linked glycan structures from hCG isolated from urine collected from choriocarcinoma patients plus several unique glycans that were not seen in hCG collected from healthy pregnancy urine (Chapter 4). The structures and compositions of N-glycans detached from hCG were compared between healthy and pathological samples. There was neither a significant difference in the overall number of glycans detected between these groups nor in the size distribution. However, within normal pregnancy urine, there was a decrease in the proportion of single antenna glycans and an increase in multi-antenna glycans (>1antenna) with increasing gestation. Also, as the pregnancy progressed, there was a drop in the proportion of non-fucosylated glycans and a concomitant increase in the proportion of multi fucosylated hCG N-linked glycans, a steady

decrease in the number of sialic acid groups attached, as well as an increase in the proportion of bisected glycans.

Our statistically significant findings showed that analysis of hCG glycosylation patterns and detection of the disorder-specific library of N-linked glycans might be used to identify choriocarcinoma from a non-pathological healthy pregnancy.

5.2 Glycans in pregnancy

Glycosylation has been among the most important post-translation modifications of glycoproteins and contributes to the instrumental role of glycans in a normal pregnancy. As Trowsdale and Betz (2006) speculated the alterations in the glycosylation of serum proteins such immunoglobulin G (IgG) and 1-antitrypsin (AAT) might be linked to the partially suppressed maternal immune system which is required for healthy fetal development during early pregnancy (Szekeres-Barthó, 2002, Ruhaak *et al.* 2014). In support of this hypothesis, it was shown that the increase in sialylation and galactosylation of the IgG fragment crystallisable (Fc) domain associated with the progression of pregnancy triggered the conversion of IgG role into an anti-inflammatory agent (Orczyk-Pawiłowicz *et al*., 2012). Moreover N-linked glycosylation and high sialic acid content influence protein folding, subunit assembly and receptor binding activity. Thus, the formation of bisecting glycans and N-glycan branching with high-mannose and complex glycans found on various proteins are often displayed during a healthy pregnancy (Varki *et al.,* 2017).

Our data revealed a highly heterogeneous repertoire of N-linked glycans in normal pregnancy. There were changes in structure with advancing gestation such as a decrease in non-fucosylated glycans, a concomitant increase in the proportion of multi fucosylated hCG N-linked glycans, and variation of sialic acid number and an increase in the proportion of bisected glycans with advancing gestation was also detected.

Over the years, several groups described fluctuations and abundance of N-linked glycans observed in normal pregnancy (Cole, 1987, Trowsdale and Betz, 2006, Stanley *et al.,* 2009). These multiple site-specific N-linked glycans with a variable number of antennas and various levels of sialylation and/or fucosylation among other structural features have been detected. Hence, the term "hyperglycosylated hCG" (hCGh) was used in the early 1990s to describe a collection of these distinct glycan groupings mentioned above largely found in early pregnancy and trophoblastic disease (Nemansky *et al*., 1995, Cole *et al.,* 1999, Gervais *et al.,* 2003). Further research by Jacoby, *et al.,* (2000) confirmed and also determined specific nine bi- and tri-antennary glycoforms between 876.8Da - 2598.4Da contributing to the microheterogeneity

of hCGβcf purified from pregnancy urine using MALDI-TOF MS. Other examples of hCG N-glycans libraries of various structural sizes ranging from 2149 Da to 5859.4 Da (Valmu *et al.,* 2006), 1421 Da to 2207 Da (Toll *et al.,* 2006), 1257.4-3620 Da (Jansen *et al.*, 2016), or most recently 1579-3923 Da (Ibeto *et al.,* 2020) were also revealed and therefore a myriad of all N- and O-linked hCG structures found in normal pregnancy have been defined as hyperglycosylated hCG. In addition, Black and Bowers, (2010) reported biantennary, fucosylated and sialylated glycans of hCG β-subunit detected using LC-MS. There have also been several attempts to characterise hCG glycoforms from commercial preparations of Pregnyl (urinary hCG) and Ovitrelle (recombinant hCG). They are both used as a fertility treatment for stimulation of ovaries, followed by ovulation, and therefore their composition is closely matching hCG structures found in pregnancy (Koistinen *et al.*, 2019). Camperi *et al.,* (2018) found that αhCG glycoforms were presented with six N-linked hCG glycans from Ovitrelle and five on Pregnyl. More detailed studies by our Biomarkers Group at Middlesex University (2021) presented the further discovery of five N-linked glycans on βhCG (1638Da-2368Da) on Ovitrelle including biantennary, fucosylated and sialylated glycoforms. Additionally a mixture of five bi-and tetra-antennary fucosylated and sialylated glycoforms (1786Da-2368Da) on Pregnyl (Ward *et al.,* 2021, awaiting publication).

Therefore, by precisely characterising hCG N-glycans repertoire we could potentially determine which glycan moieties are relevant to the identification of a healthy pregnancy and which might serve as potential disease biomarkers.

5.3 Glycans in choriocarcinoma

In urine from choriocarcinoma patients, fourteen N-glycan structures specific to choriocarcinoma were discovered which were not seen in normal pregnancy urine. In addition, there were several significant differences in the structures seen in choriocarcinoma urine.

5.3.1 Fucosylation

The proportion of multi-fucosylated glycans was three times higher in choriocarcinoma urine. We are not aware of any previous publication that has specifically associated increased hCG fucosylation in association with choriocarcinoma or any other GTN and so our finding here is highly original. Previous reports have identified increased fucosylation in association with malignant cell transformation in other cancers. It has been known for many years that hepatocellular carcinoma is associated with an increase in fucosylated alpha-fetoprotein (AFP) (Aoyagi *et al.,* 1985, 1986). The cancer antigen CA 19-9 has also been used for many years as

a tumour marker most notably for pancreatic cancer and is also fucosylated (Szymendera *et al.,* 1986). More recently, fucosylated haptoglobin has been reported to be a useful marker for the detection of pancreatic cancer (Okuyama *et al.,* 2006). Fucosylated N-linked glycans collected for ovarian tissue samples have also been reported as discriminators of ovarian cancers (Everst-Hass *et al.,* 2016). Fucosylation is also detected on prostate-specific antigen (PSA) used as a marker for prostate cancer (Butler and Huang 2021). A global plasma glycan screening approach revealed an increase in specific core-fucosylated glycans (Doherty *et al.,* 2018).

A potential role for fucosylation is indicated from several studies using α1,6-fucosyltransferase (*FUT8*)-knock-out mice. As previously mentioned *FUT8* is unique in that it is the only *FUT* responsible for core-fucosylation on N-glycoproteins (Schneider *et al.,* 2017). Core-fucosylation is required for the binding of the EGF to receptor EGFR (Wang *et al.,* 2006). Deficient migration of embryonic fibroblasts (Zhao *et al*., 2006). Since the removal of this protein greatly suppresses cancer cell functions such as proliferation and migration, it seems possible that its overexpression and consequent increase in fucosylation may enhance these pro-cancer properties. In support of this idea are the findings of Hoti *et al.,* 2018 who demonstrated a strong correlation between *FUT8* expression and aggressive prostate cancer. Furthermore, they were able to show that *FUT8* overexpression was associated with the development of androgen-resistant prostate cancer. In addition, *FUT8* is upregulated in aggressive breast carcinoma and thyroid cancer (Ito *et al.,* 2003, Tu *et al.,* 2017).

5.3.1.1 Glycan extensions: Lewis^X (Le^x) and type-2 glycan units (LacNAc)

Furthermore, fucosylated glycans synthesized by fucosyltransferases (Fuc-Ts) encoded by *FUT3* (also known as Lewis gene, *Le*) can be divided into terminal fucosylation (such as Le^x and sialyl sLe^x) and core fucosylation (Phino and Reis, 2015). Croci *et al.,* (2014) and Dimitroff (2019) discovered the synthesis of fucosylated sialyl Lewis (sLe^X) glycan epitopes, with the most generally recognised ability of sLe^x identified as stimulating metastasis and promoting tumour cell survival by binding to a family of endothelial (E) and platelet (P) selectins expressed on endothelial cells after stimulation by IL-1 and TNF- α (Croci *et al.,* 2014, Dimitroff, 2019).

In this study, terminal and core fucosylation were detected in urine from choriocarcinoma patients, which is commonly associated with metastatic cancers (Gomes *et al.,* 2013) and previously mentioned in the above section as carbohydrate antigen 19-9 (CA19-9) is a sialylated Lewis antigen of MUC1 protein found in patients with pancreatic adenocarcinoma (Cen, *et al.,* 2014) indorsing our findings. Furthermore, the enhanced expression of sLe^x was

previously described in leukaemia cells (*FUT7*) by Hiraiwa *et al.,* (2003), breast cancer (*FUT6*) by Matsuura *et al.,* (1998) or chronic bronchitis by Colomb *et al.,* 2012 among many others.

Other possible mechanisms, closely linked to this study, by which increased N-linked glycans extension and branching, enhance the progression of cancer, include lattice formation via galectin binding to (poly)-N-acetyllactosamines (LacNAc) [Galβ1–4GlcNAcβ1–3]n (of variable length) (Stanley and Cummings, 2017). These structures can act as backbone polymers for Lewis antigen structures and their altered extensions have previously been described to regulate cell adhesion processes during cancer such as hepatocellular carcinoma, colon, colorectal and lung cancer as recently reviewed by Kadirvelraj *et al*., (2021). Synthesis of LacNAc requires two enzymes; the principal poly-LacNAc synthase i.e. β1,3-GlcNAc transferases (*B3GNT2*) and β1,4-Gal transferases (*B4GALT*). This observation was supported by *B3GNT2*-knockout mice displaying a significant reduction in poly-LacNAc structures (Henion and Schwarting, 2014).

5.3.2 Sialyation

We found an increase in the proportion of multiple core- or antennal sialylations (two, three, or four sialic acid residues) in choriocarcinoma urine compared to that from first-trimester normal pregnancy. Increased sialylation is associated with the progression of many cancers. This is believed to be mainly due to increased expression of the genes corresponding to the sialyltransferases which are responsible for this modification. Silyltransferase expression is increased in choriocarcinoma along with gastric, breast, colorectal, cervical cancers and acute myeloid leukaemia (reviewed by Rodriguez and Macauley 2018).

This association is supported by numerous knockout and overexpression studies (reviewed by Dobie and Skropeta 2021). For example, a specific knockout of the sialyltransferase *ST3Gal III* gene in melanoma cells reduced α-2,3-sialylation from melanoma cells and also decreased their metastatic characteristics (Singh and Choi 2019). Furthermore, in a mouse colon cancer model, it was found that overexpression of sialyltransferase *ST8SiaVI* resulted in increased tumour growth and significantly decreased survival time (Friedman *et al*., 2019).

Several proposed mechanisms are showing how sialylation directly affects tumour progression. Similar to the role of fucosylation above, it is suggested that sialylation activates growth factor-receptor and cluster differentiation signalling pathway. For example, in ovarian cancer cells, α-2,6-sialylation of fibroblast growth factor receptor 1 (FGFR1) activates the extracellular signal-regulated protein kinase (ERK) and focal adhesion kinase (FAK) pathways, which in turn promote cell proliferation and migration (Ou *et al.,* 2020). In breast cancer, silencing of

sialyltransferase II (*ST6Gal II*) resulted in inhibition of cell cycle progression as well as reduced expression of markers of cancer migration and metastasis and patient survival (Cheng *et al*., 2020). These activities of sialylated glycoproteins are attributed to them acting as the binding ligands for key ligands involved in metastasis and tumour progression. Increased sialylation may also support tumour development through inhibition of apoptosis through inactivation of Fas ligand (Swindall *et al.,* 2011). Hypersialylation on cancer cells also helps the cancer cells evade the host immune system by binding sialic acid-binding immunoglobulin-type lectins (Siglecs), which are found on the surface of immune cells. Once bound, siglecs promote immunosuppressive signalling, thus conferring protection on the tumour (Van der Wall *et al*., 2020).

With regard to hCG, previous studies indicate that changes in sialylation have a considerable effect on its biological activity. For example, decreasing sialylation reduces the affinity of hCG for the LH receptor (O'Connor *et al.,* 1994). Nemansky and co-authors (1995) conducted a series of studies in which hCG was selectively desialylated. They found that desialylation decreased hCG bioactivity which was, in turn, restored following alpha 6-sialylation. In addition, Kovalevskaya *et al.,* 2002 and Nishino, *et al.,* 2017 defined hCG sialylation as a step leading to the high expression of tumour-associated antigens, which promotes cell detachment from the tumour mass. This promotes the adhesion of tumour cells to vascular endothelial cells and the formation of cancer metastases. Furthermore, Cole, (1998) and Handschuh *et al.,* (2007) described changes in hCG glycosylation as to be associated with invasive trophoblast cell function and Lundin *et al*., (2001) linked them to detection of ovarian and cervical malignancies, bladder (Mora *et al*., 1996), and lung cancer (Yokotani *et al.,* 1997), Down syndrome pregnancy (Cole *et al*., 1999) and other pregnancy complications or pregnancy loss (Kovalevskaya *et al.*, 2002).

5.3.3 Branching/antennas/size

Glycans composition in choriocarcinoma revealed a considerable increase in the proportion of higher molecular weight (>2100Da) glycans compared to glycans found in the first trimester of a healthy pregnancy. This was consistent with a 30% increase in tri- and tetrantennary and a 10% drop in mono- and di-antennary structures. Cancer has been associated with changes in the degree of N-glycans' branching which closely correlates with the size of N-glycans. Accordingly, enzymatic activities of α-mannosyl- β1,6 N-acetylglucosaminyltransferase-V (GnT-V; *MGAT5*) generate an increase in N-glycan branching (tri- and tetra- antennary) that can modify a protein's half-life, stability, membrane dynamics, extracellular-binding partners,

and functional activity and have been detected in many cancers (Dimitroff, 2019). This observation was previously presented by Seberger and Chaney, (1999) who found that *MGAT5* expression promotes adhesion, migratory activity and metastasis during *in vitro* study of breast and lung cancer. Similarly, reviewed by Moriwaki and Miyoshi, (2010) and Pinho *et al.,* (2013) enhanced branching of N-glycans during prostate and colon cancers caused by dysfunction of N-acetylglucosaminyltransferase-III and V (GnT- III; *MGAT 3* and GnT-V) was observed and described as a hallmark of invasive carcinomas. The variations of GnT-III which regulates E-cadherin-mediated tumour suppression and GnT-V which causes E-cadherin-mediated tumour invasion were reported as bidirectional crosstalk between E-cadherins in cancer. The effect of branched N-glycans was also discussed by Link-Lenczowski *et al.,* (2018) suggesting that melanoma progression mainly occurs through the modulation of integrin-dependent adhesion and migration which is directly related to the overexpression of GnT-III. Also, downregulation of *MGAT3* in the mouse model of breast cancer cells (PyMT) induced migratory properties of cancer cells promoting the metastases to the lung (Song *et al.*, 2010).

Alternative choriocarcinoma *in vitro* study conducted by Fukushima *et al.* (1999) on JEG-3 (Homo sapiens placenta choriocarcinoma) and BeWo (Homo sapiens placenta choriocarcinoma) cell lines found that N-acetylglucosaminyltransferase IV (GnT-IV) expression, was significantly higher generating abnormal biantennary structures in choriocarcinoma than in the normal trophoblast. Following this observation Niimi, *et al*., (2012) established that GnT-IVa was involved in regulating the invasion of choriocarcinoma. Further *in vivo* studies on mice confirmed that knockdown of GnT-IVa suppressed cell adhesion, migration, tumour engraftment and growth (Niimi *et al.,* 2012). And five years later after additional research on Jar (Homo sapiens placenta choriocarcinoma) cell lines Nishino *et al*., (2017) concluded that GnT-IVa promoted and regulated the invasion of choriocarcinoma through modifications of the carbohydrates chains.

5.3.4 Bisection and Sulphation

Bisection is a unique form of N-glycosylated alteration consisting of an additional β1,4-linked GlcNAc bound to the core β-mannose catalysed by glycosyltransferases i.e. GnT-II, GnT-IV and GnT-V (*MGAT 3*) (Miwa *et al.,* 2012, Lu *et al*., 2016). On the other hand, the synthesis of sulphated N-linked glycans is controlled by β1-4 GalNAcT either encoded by *B4GALNT3* or *B4GALNT4.* The terminal β1-4 GalNAc is sulphated by sulphotransferase *CHST8* or *CHST11* which are expressed in pituitary cells. A sulphate group in complex N-linked glycans is commonly localized on C-3 of galactose (Gal), C-6 of N-acetylglucosamine (GlcNAc) or C-

4 of N-acetylgalactosamine (GalNAc) of terminal Gal-GlcNAc and GalNAc-GlcNAc residues matching our current findings for hCG glycoprotein presented in Table 4.4 (Chapter 4) (She *et al*., 2019).

In this study, we found a significant increase in the proportion of bisected and sulphated hCG N-glycans detected in the urine of choriocarcinoma patients compared to a first-trimester healthy pregnancy. Bisected glycans have been linked to a variety of biological processes, including cell division and adhesion, fertilisation and fetal development, and tumour growth (Chen *et al.,* 2020). In cancer, GnT-V were linked to tumour cell metastasis and carcinogenesis (Gu *et al*., 2009) and bisecting glycan structures were identified in ovarian cancer (Sethi, *et al.,* 2014) and breast cancer (Tan *et al.,* 2018). *MGAT3* altered epithelial growth factor receptor (EGFR) and integrin N-glycosylation and reduced EGFR/integrin cell surface binding to galectin-3, resulting in their endocytosis, inhibition of intracellular signalling, and encouragement of cell motility and tumour spread (Partridge *et al.,* 2004). Furthermore, Ibeto *et al.* (2020) characterised the presence of bisected type N-glycans on hCG detected in gestational trophoblastic diseases (GTD) supported previous findings of differential glycosylation of trophoblast cells by Chen and co-workers (2016). It was suggested that the expression of elevated levels of bisected glycans detected in the layer of syncytiotrophoblasts and cytotrophoblasts found in human placental villi cells suppress natural killer cell-mediated cytotoxicity *in vitro* (King and Loke, 1990, Chen *et al.,* 2016).

Furthermore, negatively charged-sulphated glycans were also recently recognised and identified as derivatives from pituitary hormones such as luteinizing hormone (LH) and thyroid-stimulating hormone (TSH) but not follicle-stimulating hormone (FSH) (Cahoreau, *et al.,* 2015). Additionally, sulphated glycans were used to distinguish changes in gene expression linked with sulphation between two main types of lung cancer: small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). Toyoda *et al.,* (2016) discovered that SCLC cell lines almost exclusively expressed two genes, *CHST8* and *CHST9*, which encode the sulfotransferase isozymes GalNAc4ST1 and GalNAc4ST2, respectively. These sulfotransferases are important enzymes in the manufacture of sulfated glycoproteins, some of which are released into the bloodstream and used as biomarkers for SCLC. In addition, Horlacher *et al.* (2010) investigated galectin sugar-binding preferences and found that galectin-1 and galectin-2 preferentially bind sulfated glycans. Thus, galectin-1 was shown to be upregulated in colorectal cancer and to correlate with the degree of dysplasia, suggesting that galectin-1 is linked to colorectal cancer development (Horlacher *et al.,* 2010).

Sulphated hCG was also considered to promote androstenedione synthesis, as well as ovulation and corpus luteum development, during the follicular phase of the cycle. It may also help boost progesterone production during the luteal phase (Cole and Gutierrez, 2009, d'Hauterive *et al.,* 2022).

Besides, mentioned above functions of bisected and sulphated glycans in cancer, glycan bisection plays also an important role in neurological systems. Hence, studies on the mouse brain stem by Shimizu *et al.,* (1993), Shigeta *et al.,* (2006) and more recently by Nagae *et al.,* (2016) detected bisected glycans which promoted β1 integrin-mediated neuritogenesis induced by GlcNAc-T III. This association was connected to progressive, neurodegenerative Alzheimer's disease where glycoproteins such as tau, Ab-precursor protein (APP), and b-site APP-cleaving enzyme-1 (BACE-1) (Schedin-Weiss *et al.,* 2014, 2019; Kizuka *et al.,* 2015). Moreover, sulphated glycans were also recognised in numerous immune system dysfunctions such as cystic fibrosis (Chance and Mawhinney, 1996) however, the pioneering discovery by Wang *et al.,* (2017) allowed to identify sulphated glycans on IgGs as a biomarker for rheumatoid arthritis (Wang *et al.,* 2017). The extensive glycosylation of ACPA-IgG variable domains was found to modulate binding to citrullinated antigens and demonstrated a pro-inflammatory glycosylation phenotype (Rombouts *et al.,* 2016).

In conclusion, it is still not clear which of the described above protein glycosylation changes have the main role in the pathogenesis of choriocarcinoma however, tumorigenesis and cancer progression is supported by findings of Kovalevskaya *et al.,* (2002) and Nishino *et al.,* (2017) who showed that hCG carbohydrate moieties are very important tumour markers as it activates invasion of choriocarcinoma and germ cells which might contribute to malignancy. Glycans can directly bind promalignant glycan-binding proteins and engage in the activation of growth factor receptors, integrins, integrin ligands, lectins, and other type-1 transmembrane proteins. Moreover, discussed above group of post-translation modifications leads to the high expression of tumour-associated antigens, which promotes cell detachment from the tumour mass and inhibits cell-cell adhesion. This promotes the adhesion of tumour cells to vascular endothelial cells and the formation of cancer metastases. Although the biological activity of hCG diminishes with the gradual desialylation or partial, or complete removal of carbohydrate units internal to the sialic acid, the removal of certain residues of carbohydrates increases the affinity of hCG for its receptor (O'Connor *et al.,* 1994).

Thus, altered N-glycans post-translation modifications found on glycoproteins have been seen in several cancer types, which adds to the instability and activity of the target protein and is associated with invasiveness and poor prognosis in cancer patients (Pietrobono and Stecca, 2021).

Consequently, high precision characterisation of hCG glycoforms using MALDI-TOF could enable us to demonstrate differences in the composition and structure of N-linked glycans found in healthy pregnancy and urine collected from choriocarcinoma patients. Additionally, concentration, structure and glycosylation occupancy on a certain site of glycosylated protein may simultaneously vary, providing distinct information during the progression of the disease (Zhang *et al.,* 2014). As a result, not only quantitative data but also information on glycosylation forms, including changes in glycan composition and structure, are significant in glycomic research, as we propose in this study.

5.4 Limitations in detection

Currently, immunoassays are used for hCG detection based on their expression under diverse clinical situations such as healthy pregnancy but also for detection of ectopic pregnancies, management of hyperemesis gravidarum, diagnosis and management of trophoblastic diseases and testicular, and germ cell neoplasms (Cole *et al.,* 2001). As a result, the selectivity of immunoassays for individual hCG isoforms including intact hCG (α+β subunits), free β subunit, nicked hCG and hyperglycosylated hCG has been restricted and determined by the selectivity of antibodies employed. Consequently, detecting a wide variation in the structure of the hCG molecule has proven to be problematic, yielding false-positive or false-negative findings (Butler *et al.,* 2001). Furthermore, persistent mild elevation of hCG, caused by pathological conditions still not fully understood; such as familial hCG (found in both men and women) or phantom hCG and phantom choriocarcinoma, where hCG concentration is constantly elevated (10-200 IU/L in serum and urine sample) has been more frequently diagnosed and reported (Cole, 1998, Cole, 2012, Tan *et al*., 2014). Currently, tests cannot distinguish between specific hCG isoforms, which leads to patient treatment with cytotoxic chemotherapy when in fact, no true hCG or GTD is present. False-positive hCG assays results may even lead to unnecessary surgery and chemotherapy (Cole, 1998, Cole *et al*., 1999, ACOG, 2002, Soares *et al.,* 2016). In addition, McCash *et al.,* (2017) presented that the false positive tests (based on current antibodies) cause misdiagnosis of 38% of patients, which might have a huge impact on hospitals' budgets and therefore a need for a new independent approach for hCG isoforms is required (McCash *et al.,* 2017, Gao *et al.,* 2019).

Analysis of glycosylation remains a challenge as N-linked glycans have low ionization efficiency, labile nature and complicated branched structures involving various linkage isomers as described above. Thus, MALDI-TOF MS is often used as a first step for N-linked structural glycan profiling due to its low sample consumption, high analysis speed and high automation potential. Results of high peak intensities and good baseline separation can be achieved by performing optimized methods for sample application on the target plate, use of the appropriate sample to matrix ratio and sufficient sample purity as described in this study. However, we are also aware of the limitations in our method linked to the lack of differentiation between isomeric species (compounds with identical mass and composition) and the small sample size. Also, the presence of sialic acid residues on N-glycans can cause a decrease in ionization efficiency and instability of sample in- and post-source decay however, the charge state of sialylated compounds is also surprisingly higher (Ruhaak *et al.*, 2009; Nishikaze, 2017). However, MALDI-TOF is a semi-quantitative technique, so quantification studies would require additional derivatization steps.

5.5 Summary

In the course of this research, we have developed a rapid mass spectrometry-based method for the identification and detection of carbohydrate moieties on human chorionic gonadotropin by combining the deglycosylation procedure for extraction and purification of N-linked glycans and implementation of the MALDI-TOF mass spectrometry with GlycoQuest software. In conclusion, we have shown that the recuperation of hCG from clinical urine using the Dyneabeads purification method was very successful (87% recovery). 1mL of urine was sufficient to establish heterogeneity of hCG glycoforms present in physiological and pathological disorders during pregnancy. Further, a characteristic glycan profile of a normal pregnancy and choriocarcinoma was determined and could lead to the development of new, diagnostic approaches used in the identification of various diseases.

5.6 Implications for future research

The primary benefit of mass spectrometry is the strength of the analytical evidence connected to the absolute identification of the target analytes which is a driving force in the personalized approach to disease diagnosis and treatment. Therefore, reviewing clinical conditions for example preeclampsia, hyperemesis gravidarum, ectopic pregnancies, diabetic pregnancies or gestational trophoblastic diseases and finding a biochemical marker using MALDI-TOF MS would help with early detection and quick management of the condition. Changes in glycan

structures and composition could be used as new, more sensitive and specific glyco-biomarkers. We can also confidently suggest that combining proteomics, glycomics, and genomics studies could lead to further discoveries in the hCG field.

Also, to fully assign all glycan structures a combination of analyses will be required, including MS/MS analysis to distinguish between different isomeric structures. The MS/MS CID (collision-induced dissociation) and HCD (higher-energy collisional dissociation) help with linkage position assignment (site-specific N-glycosylation). Enzymatic treatment to further characterise glycosyltransferase specificity. Additionally, permethylation for the derivatisation of glycans for MALDI-MS detection and quantification (procainamide, reductive amination (2-AB, 2-AA) or permethylation Kit with Methyl Iodide (MeI) (Ludger, UK) could be applied. Finally, consideration of a larger sample size will help with a more detailed structural characterisation of glycans.

In conclusion, the ultimate goal will be the development of site-specific methods capable of simultaneously characterising and quantifying multiple glycoprotein isoforms with multiple glycan configurations attached at various glycosylation sites with minimal sample preparation.

References

Abbott KL and Pierce JM (2010) Lectin-based glycoproteomic techniques for the enrichment and identification of potential biomarkers. Methods Enzymol. 480:461-76

Alldred SK, Guo B, Takwoingi Y, Pennant M, Wisniewski S, Deeks JJ, Neilson JP, Alfirevic Z (2015) Urine tests for Down's syndrome screening. The Cochrane database of systematic reviews, 2015(12)

American Cancer Society (2017) https://www.cancer.org/cancer/gestational-trophoblastic-disease/about/what-is-gtd.html

Aoyagi Y, Isemura M, Suzuki Y, Sekine C, Soga K, Ozaki T, Ichida F (1985) Fucosylated alpha-fetoprotein as marker of early hepatocellular carcinoma. Lancet, 2: 1353-1354

Aoyagi Y, Isemura M, Suzuki Y, Sekine C, Soga K, Ozaki T, Ichida F (1986) Change in fucosylation of alpha-fetoprotein on malignant transformation of liver cells. Lancet, 1: 210

Arabi A, Ayoola-Adeola M, Nguyen HQ, Brar H, Walker C (2021) Pulmonary metastasis as a primary manifestation of gestational choriocarcinoma in a third trimester pregnancy. Gynecol Oncol Rep. 31;36:100762

Armbruster DA, and Pry T (2008) Limit of blank, limit of detection and limit of quantitation. The Clinical biochemist. Reviews, 29 Suppl 1(Suppl 1), S49–S52

Azevedo R, Silva A, Reis CA, Santos LL, Ferreira JA (2018) In silico approaches for unveiling novel glycobiomarkers in cancer, Journal of Proteomics 171, p. 95–106

Bapiro TE, Richards FM, Jodrell DI (2016) Understanding the Complexity of Porous Graphitic Carbon (PGC) Chromatography: Modulation of Mobile-Stationary Phase Interactions Overcomes Loss of Retention and Reduces Variability, Analytical Chemistry 88 (12), 6190-6194

Bapiro TE, Richards FM, Jodrell DI (2016) Understanding the Complexity of Porous Graphitic Carbon (PGC) Chromatography: Modulation of Mobile-Stationary Phase Interactions Overcomes Loss of Retention and Reduces Variability, American Chemical Society, 88, p.6190-6194

Berg JM, Tymoczko JL, Stryer L, Stryer L (2002) Biochemistry. New York: 18th edition, W.H. Freeman.

Berger P and Lapthorn AJ (2016) Standardization of Epitopes for Human Chorionic Gonadotropin (hCG) Immunoassays Current Medicinal Chemistry 23:(999)

Berger P, Paus E, Hemken PM, Sturgeon C, Stewart WW, Skinner JP, Harwick LC, Saldana SC, Ramsay CS, Rupprecht KR, Olsen KH, Bidart JM, Stenman UH (2013) Candidate epitopes for measurement of hCG and related molecules: the second ISOBM TD-7 workshop, Tumour Biol. 34(6): p.4033-57

Berger P, Sturgeon C, Bidart JM, Paus E, Gerth R, Niang M, et al. (2002) The ISOBM-TD-7 workshop on hCG and related molecules. Towards user-oriented standardization of pregnancy and tumor diagnosis:assignment of epitopes to the three-dimensional structure of diagnostically and commercially relevant monoclonal antibodies directed against human chorionic gonadotropin and derivatives. Tumor Biol, 23:1–38

Birken S, Gawinowicz MA, Kardana A, Cole LA (1991) The Heterogeneity of Human Chorionic Gonadotropin (hCG). II. Characteristics and Origins of Nicks in hCG Reference Standards, Endocrinology, Volume 129, Issue 3, 1, 1551–1558,

Blithe DL, Richards RG, Skarulis MC (1991) Free alpha molecules from pregnancy stimulate secretion of prolactin from human decidual cells: a novel function for free alpha in pregnancy. Endocrinology 129, 2257–2259

Blithe, DL (1990) Carbohydrate composition of the alpha-subunit of human choriogonadotropin (hCG alpha) and the free alpha molecules produced in pregnancy: most free alpha and some combined hCG alpha molecules are fucosylated. Endocrinology, 126(6) 2788-2799

Bolstad N and Warren DJ (2013) Heterophilic antibody interference in immunometric assays Best Practice & Research Clinical Endocrinology & Metabolism 27 (5), 647-661

Braunstein GD, Kamdar V, Rasor J, Swaninathan N, Wade ME (1979) Widespread distribution of a chorionic gonadotropin-like substance in normal human tissues. Clin. Endocrinol. Metab., 49, 917- 925

Brockhausen I (2006), Mucin-type O-glycans in human colon and breast cancer: glycodynamics and functions, EMBO Rep. 7(6), 599–604

Brown J, Naumann RW, Seckl MJ, Schink J (2017) 15 years of progress in gestational trophoblastic disease: Scoring, standardization, and salvage, Gynecologic Oncology, 144, 200-207

Bulun SE (2011) Physiology and pathology of the female reproductive axis. In: Melmed S, Polonsky KS, Larsen PR, Kronenberg HM (Eds.), Williams Textbook of Endocrinology, 12th ed. Elsevier Saunders, Philadelphia, PA, 581–660

Burczynska B, Booth MJ, Iles RK, Shah A, Shiled A, Butler S (2013) Stable knockdown of hCGβ mRNA expression in bladder cell results in significant growth inhibition, Anticancer Research, 33, 1-6

Buskas T, Ingale S, Boons G (2006) Glycopeptides as versatile tools for glycobiology, Glycobiology, 16(8), 113R–136R

Butler SA, Khanlian SA, Cole LA (2001) Detection of early pregnancy forms of human chorionic gonadotropin by home pregnancy test devices. Clin Chem, 47(12):2131-6

Butler W and Huang J (2021) Glycosylation Changes in Prostate Cancer Progression , Frontiers in Oncology, 11: 5502

Cahoreau C, Klett D, Combarnous Y (2015) Structure-function relationships of glycoprotein hormones and their subunits' ancestors Front Endocrinol (Lausanne) 26;6: 26

Cao J, Chen Y, Yang P (2009) Analytical progress for protein glycosylation in China. Front. Chem. China 4, 360

Carvalho S, Oliveira T, Bartels MF, Miyoshi E, Pierce M, Taniguchi N, Carneiro F, Seruca R, Reis CA, Strahl S, Pinho SS (2016) O-mannosylation and N-glycosylation: two coordinated mechanisms regulating the tumour suppressor functions of E-cadherin in cancer. Oncotarget, 7(40), 65231–65246

Cen P, Walther C, Finkel KW, Amato RA (2014) Chapter 3 - Biomarkers in Oncology and Nephrology, Renal Disease in Cancer Patients, Academic Press, 21-38

Chance DL and Mawhinney TP (1996) Disulfated Oligosaccharides. Derived from Tracheobronchial Mucous Glycoproteins of a Patient Suffering from Cystic Fibrosis. Carbohydr. Res. 295, 157–177

Chen F, Wang Y, Puett D (1992) The carboxy-terminal region of the glycoprotein hormone alpha-subunit: contributions to receptor binding and signaling in human chorionic gonadotropin. Mol Endocrinol. 6(6): 914-9

Chen Q, Tan Z, Guan F, Ren Y (2020) The Essential Functions and Detection of Bisecting GlcNAc in Cell Biology, Frontiers in Chemistry, 8

Cheng J, Wang R, Zhong G, Chen X, Cheng Y, Li W et al. (2020) ST6GAL2 Downregulation inhibits cell adhesion and invasion and is associated with improved patient survival in breast cancer. Onco.Targets Ther. 13, 903–914

Cheung A, Juan Zhang H, Chen Xue W, Siu M (2009) Pathogenesis of choriocarcinoma: Clinical, genetic and stem cell perspectives. Future Oncology 5(2):217-31

Cocquebert M, Berndt S, Segond N, Guibourdenche J, Murthi P, Aldaz-Carroll L, Evain-Brion D, Fournier T. (2012) Comparative expression of hCG β-genes in human trophoblast from early and late first-trimester placentas. Am J Physiol Endocrinol Metab, 15 (8) 303

Cole LA (1987) The O-linked oligosaccharide structures are striking different on pregnancy and choriocarcinoma HCG. J Clin Endocrinol Metab. 65(4):811-3

Cole LA (1997) Immunoassays of human chorionic gonadotropin, its free subunits and metabolites, Clin, Chem, 43, 2233-2243

Cole LA (1998) Phantom hCG and phantom choriocarcinoma. Gynecol Oncol. 71(2) 325-9

Cole LA (2007) Hyperglycosylation hCG, Placenta, 28, 977-986

Cole LA (2009) Human chorionic gonadotropin and associated molecules. Expert Rev Mol Diagn. 9(1) 51-73

Cole LA (2009) New discoveries on the biology and detection of human chorionic gonadotropin, Reprod Biol and Endocrinol, 7(8) 1-37

Cole LA (2010) Biological functions of hCG and hCG-related molecules. Reprod Biol Endocrinol. 8:102

Cole LA (2010) Structurally related molecules of HCG in GTD, Elsevier, USA,148-182

Cole LA (2012) Familial hCG Syndrome, J. Rep Immuno 93 (1), 52-57

Cole LA (2012) Hyperglycosylated hCG and pregnancy failures. Journal of Reproductive Imm. 93(2):119-122

Cole LA and Gutierrez JM (2009) Production of human chorionic gonadotropin during the normal menstrual cycle. J. Reprod. Med. 54, 245–250

Cole LA, Khanlian SA, Riley JM, Butler SA (2006) Hyperglycosylated hCG in gestional implantation and in choriocarcinoma and testicular germ cell malignancy tumorigenesis, J.Reprod Med, 51, 919-925

Cole LA, Rinne KM, Mahajan SM, Oz UA, Shahabi S, Mahoney MJ, Bahado-Singh RO (1999) Urinary screening tests for fetal Down syndrome: I. Fresh beta-core fragment. Prenat Diagn. 19(4) 340-50

Cole LA, Shahabi S, Oz UA, Bahado-Singh RO, Mahoney MJ (1999) Hyperglycosylated human chorionic gonadotropin (invasive trophoblast antigen) immunoassay: A new basis for gestational Down syndrome screening. Clin Chem, 45(12): 2109-2119

Colomb F, Krzewinski-Recchi M, El Machhour F et al, (2012) TNF regulates sialyl-Lewisx and 6-sulfo-sialyl-Lewisx expression in human lung through up-regulation of ST3GAL4 transcript isoform BX, Biochimie 94, 2045-2053

Cooper JA (2016) https://radiologykey.com/gestational-trophoblastic-neoplasia/

Crum CP (2005) The female genital tract. In Kumar V, Robbins SL, Cotran RS [eds]: Pathologic Basis of Disease, 7th ed. Philadelphia, WB Saunders, An Imprint of Elsevier

d'Hauterive SP, Close R, Gridelet V, Mawet M, Nisolle M, Geenen V (2022) Human Chorionic Gonadotropin and Early Embryogenesis: Review" International Journal of Molecular Sciences 23, 3: 1380

de Hoffmann E and Stroobant V (2007) Mass Spectrometry: Principles and Applications, 3rd Edition

de Medeiros SFand Norman RJ (2009) Human choriogonadotrophin protein core and sugar branches heterogeneity: basic and clinical insights, Human Reproduction Update, 15:1 69–95

Demers LM, Gabbe SG, Villee C, Greep R (1973) Human chorionicgonadotropin mediated glycogenolysis in human placenta, Biochim.Biophys.Acta, 313, 202-210

Dennis JW, Laferté S, Waghorne C, Breitman ML, Kerbel RS (1987) Beta 1-6 branching of Asn-linked oligosaccharides is directly associated with metastasis Science.1;236(4801): 582-5

Dimitroff CJ (2019) I-branched carbohydrates as emerging effectors of malignant progression PNAS, 116 (28)

Dobie C and Skropeta D (2021) Insights into the role of sialylation in cancer progression and metastasis. Br J Cancer. 124(1) 76-90

Doherty M, Theodoratou E, Walsh I, Adamczyk B, Stöckmann H, Agakov F, Timofeeva M, Trbojević-Akmačić I, Vučković F, Duffy F, McManus CA, Farrington SM, Dunlop MG, Perola M, Lauc G, Campbell H, Rudd PM (2018) Plasma N-glycans in colorectal cancer risk. Scientific reports, 8(1), 8655

Drife JO and Magowan BA, (2005) Clinical obstetrics and gynaecology, Williams&Wilkins

Duffy L, Zhang L, Sheath K, Love DR, George AM (2015) The Diagnosis of Choriocarcinoma in Molar Pregnancies: A Revised Approach in Clinical Testing. Journal of clinical medicine research, 7(12), 961–966

Elliott MM, Kardana A, Lustbader JW, et al. (1997) Carbohydrate and peptide structure of the alpha- and beta-subunits of human chorionic gonadotropin from normal and aberrant pregnancy and choriocarcinoma. Endocrine 7: 15–32

Esfandiari N, Goldberg JM (2003) Heterophile antibody blocking agent to confirm false positive serum human chorionic gonadotropin assay. Obstet Gynecol.101(5 Pt 2):1144-6

Esteves SC (2015) Efficacy, efficiency and effectiveness of gonadotropin therapy for infertility treatment Medical Express 2 (3)

Everest-Dass AV, Briggs TB, Kaur G, Oehler MK, Hoffmann P, Packer NH (2016) N-glycan MALDI Imaging Mass Spectrometry on Formalin-Fixed Paraffin-Embedded Tissue Enables the Delineation of Ovarian Cancer Tissues, Molecular & Cellular Proteomics, 15(9) 3003-3016

Fan J, Wang S, Yu S, He J, Zheng W, Zhang J (2012) N-acetylglucosaminyltransferase IVa regulates metastatic potential of mouse hepatocarcinoma cells through glycosylation of CD147. Glycoconj J. 29:323–334

Fournier, T (2016) Human chorionic gonadotropin: Different glycoforms and biological activity depending on its source of production, Annales d'Endocrinologie, 77 (2), 75-81

Frendo JL, Olivier D, Cheynet V, Blond JL, Bouton O, Vidaud M, et al. (2003) Direct involvement of HERV-W Env glycoprotein in human trophoblast cell fusion and differentiation, Mol Cell Biol, 23, 3566-3574

Friedman DJ, Shapiro M, Rajcula M, McCue S, Shapiro VS (2019) ST8Sia6 Overexpression accelerates tumor growth, alters macrophage polarization and the immune response. J. Immunol. 202, Article 135.138

Gagneux P, Aebi M, Varki A (2017) Evolution of glycan diversity In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2017. Chapter 20

Gasparri R, Sedda G, Brambilla D, Girelli L, Diotti C, Spaggiari L (2019) When a Differential Diagnosis Is Fundamental: Choriocarcinoma Mimicking Lung Carcinoma. J. Clin. Med. 2019, 8, 2018. https://doi.org/10.3390/jcm8112018

Gervais A, Hammel YA, Pelloux S, Lepage P, Baer G, Carte N, Sorokine O, Strub JM, Koerner R, Leize E, Van Dorsselaer A (2003) Glycosylation of human recombinant

gonadotrophins: characterization and batch-to-batch consistency. Glycobiology 13(3):179-89

Gillott DJ, Iles RK, Chard T (1996) The effects of beta-human chorionic gonadotrophin on the in vitro growth of bladder cancer cell lines. Br JCancer 73, 323–326

Gnoth C and Johnson S (2014) Strips of Hope: Accuracy of Home Pregnancy Tests and New Developments Geburtshilfe Frauenheilkd. 74(7):661–669

Gomes C, Osório H, Pinto MT, Campos D, Oliveira MJ, Reis CA (2013) Expression of ST3GAL4 leads to SLe(x) expression and induces c-Met activation and an invasive phenotype in gastric carcinoma cells. PloS one, 8(6)

Gondek MB, Burczynska B, Hills F (2018) Method optimisation for identification of carbohydrate moieties on human chorionic gonadotropin glycoforms using MALDI-TOF-MS. In: 45th International Society of Oncology and BioMarkers (ISOBM) 2018 Congress, 24-27 Nov 2018, Hamburg, Germany

Gosrani P (2014) Slide share, https://www.slideshare.net/poorvishah10/maldi-tof-37328025

Gross JH (2006) Mass spectrometry: a textbook. Springer Science & Business Media

Gu J, Sato Y, Kariya Y, Isaji T, Taniguchi N, Fukuda T (2009) A mutual regulation between cell-cell adhesion and N-glycosylation: implication of the bisecting GlcNAc for biological functions. J. Proteome Res. 8, 431–435

Guibourdenche J, Handschuh K, Tsatsaris V, et al. (2010) Hyperglycosylated hCG is a marker of early human trophoblast invasion. The Journal of Clinical Endocrinology & Metabolism 95(10): 240-E244

Hamid U, Royle L, Saldova R, Radcliffe CM, Harvey DJ, Storr SJ, Pardo M, Antrobus R, Chapman CJ, Zitzmann N, Robertson JF, Dwek, RA, Rudd PM (2008) A strategy to reveal potential glycan markers from serum glycoproteins associated with breast cancer progression, Glycobiology, Volume 18, Issue 12, 1105–1118

Handschuh K, Guibourdenche J, Tsatsaris V et al. (2007) Human chorionic gonadotropin expression in human trophoblasts from early placenta: Comparative study between villous and extravillous trophoblastic cells. Placenta. 28(2–3) 175-184

Harvey DJ, Wing DR, Kuster B, Wilson BH (2000) Composition of N-Linked Carbohydrates from Ovalbumin and Co-purified Glycoproteins. J Am Soc Mass Spectrom 11: 564–571

Hashim OH, Jayapalan JJ, Lee CS (2017). Lectins: an effective tool for screening of potential cancer biomarkers. PeerJ, 5, e3784

Hay DL and Lopata A (1988) Chorionic gonadotropin secretion by human embryos in vitro, J Clin Endocrinol Metab. 67(6):1322-4

Henion TR and Schwarting GA (2014) N-linked polylactosamine glycan synthesis is regulated by co-expression of beta3GnT2 and GCNT2. J. Cell Physiol. 229, 471–478

Hermsteiner M, Zoltan DR. Künzel W (2002) Human chorionic gonadotropin attenuates the vascular response to angiotensin II, Journal of Obstetrics & Gynecology and Reproductive Biology 102(2) 148-54

Hiraiwa N, et al. (2003) Transactivation of the fucosyltransferase VII gene by human T - cell leukemia virus type 1 Tax through a variant cAMP-responsive element. Blood 101, 3615–3621

Hoermann R, Spoettl G, Grossmann M et al.(1993) Molecular heterogeneity of human chorionic gonadotropin in serum and urine from patients with trophoblastic tumors. Clin Investig 71, 953–960

Hoja-Łukowicz D, Szwed S, Laidler P, Lityńska A (2018) Proteomic analysis of Tn-bearing glycoproteins from different stages of melanoma cells reveals new biomarkers. Biochimie, 151:14-26.

Holland JJ and Hreshchyshyn MM (2012) Choricarcinoma, Springer, 3rd edition

Homan G, Brown S, Moran J, Homan S, Kerin J (2000) Human chorionic gonadotropin as a predictor of outcome in assisted reproductive technology pregnancies, Rep End., 73(2): 270-274

Horlacher T, Oberli MA, Werz DB, Krouck L, Bufali S, Mishra R, Sobek J, Simons K, Hirashima M, Niki T, Seeberger PH (2010) Determination of carbohydrate-binding preferences of human galectins with carbohydrate microarrays. Chembiochem. 11, 1563–1573

Höti N, Yang S, Hu Y et al. (2018) Overexpression of α (1,6) fucosyltransferase in the development of castration-resistant prostate cancer cells. Prostate Cancer Prostatic Dis 21, 137–146

Ibeto L, Antonopoulos A, Grassi P, Pang PC, Panico M, Bobdiwala S, Al-Memar M, Davis P, Davis M, Norman Taylor J, Almeida P, Johnson MR, Harvey R, Bourne T, Seckl M, Clark G, Haslam SM, Dell A.(2020) Insights into the hyperglycosylation of human chorionic gonadotropin revealed by glycomics analysis. PLoS One. 11;15(2)

Iles R, Cole L, Butler SA (2014) Direct analysis of hCGβcf glycosylation in normal and aberrant pregnancy by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. International Journal of Molecular Sciences. 15(6) 10067-10082

Ito Y., Miyauchi A., Yoshida H., Uruno T., Nakano K., Takamura Y., Miya A., Kobayashi K., Yokozawa T., Matsuzuka F., et al. (2003) Expression of α1,6-fucosyltransferase (FUT8) in papillary carcinoma of the thyroid: Its linkage to biological aggressiveness and anaplastic transformation. Cancer Lett.;200:167–172.

Jang KS, Kim Y, Gil G, Park S, Kim B (2009) Mass spectrometric quantification of neutral and sialylated N-glycans from a recombinant therapeutic glycoprotein produced in the two Chinese hamster ovary cell lines, Analytical Biochemistry 386, 228–236

Jansen BC, Bondt A, Reiding KR, Lonardi E, de Jong CJ, Falck D, Kammeijer G, Dolhain R, Rombouts Y, Wuhrer M (2016) Pregnancy-associated serum N-glycome changes studied by high-throughput MALDI-TOF-MS, Scientific Reports 6: 23296

Jensen PH, Karlsson NG, Kolarich D, Packer NH (2012). Structural Analysis of N- and O-Glycans Released from Glycoproteins. Nat. Protoc. 7 (7), 1299–1310

Kadirvelraj R, Yang J, Kim HW, Sanders JH, Moremen KW, Wood ZA (2021) Comparison of human poly-N-acetyl-lactosamine synthase structure with GT-A fold glycosyltransferases supports a modular assembly of catalytic subsites, Journal of Biological Chemistry, 296

Khademi F, Hamzehee K, Mostafaie A, Hajihossaini R (2009) Purification of three major forms of beta-hCG from urine and production of polyclonal antibodies against them. Clin Biochem;42(13-14):1476-82

Kicman A, Parkin MC, Iles RK (2007) An introduction to mass spectrometry based proteomics-detection and characterization of gonadotropins and related molecules Mol Cell Endocrinol. 2;260-262: 212-27

King A and Loke YW (1990) Human trophoblast and JEG choriocarcinoma cells are sensitive to lysis by IL-2-stimulated decidual NK cells, Cellular Immunology, 129 (2) 435-448

Kirwan A, Utratna M, O'Dwyer ME, Joshi L, Kilcoyne (2015), Glycosylation-Based Serum Biomarkers for Cancer Diagnostics and Prognostics BioMed Research International, 15(16)

Kizuka Y, Kitazume S, Fujinawa R, Saito T, Iwata N, Saido TC. et al. (2015) An aberrant sugar modification of BACE1 blocks its lysosomal targeting in Alzheimer's disease. EMBO Mol. Med. 7, 175–189

Kliman H (2009) From Trophoblast to Human Placenta (from The Encyclopedia of Reproduction) Yale University School of Medicine

Koistinen H, Hautala L, Koli K, Stenman UH (2015) Absence of TGF-receptor activation by highly purified hCG preparations Mol Endocrinol 29(12): 1787–1791

Koistinen H, Koel M, Peters M, Rinken A, et al., (2019) Hyperglycosylated hCG activates LH/hCG-receptor with lower activity than hCG, Molecular and Cellular Endocrinology, 479, 103-109

Kovalevskaya G, Birken S, Kakuma T, et al. (2002) Differential expression of human chorionic gonadotropin (hCG) glycosylation isoforms in failing and continuing pregnancies: preliminary characterization of the hyperglycosylated hCG epitope. J Endocrinol, 172: 497–506

Kovalevskaya G, Birken S, Kakuma T, O'Connor JF (1999) Early cytotrophoblastic cells, the finding that there is a modulation pregnancy human chorionic gonadotropin (HCG) isoforms measured by an of choriocarcinoma-like HCG during early pregnancy may be immunometric assay for choriocarcinoma-like HCG. J. Endocrinol., 161, 99–106

Kovalevskaya G, Genbacev O, Fisher J, Caceres E, O'Connor JE (2002) Trophoblast origin of hCG isoforms: cytotrophoblasts are the primary source of choriocarcinoma-like hCG, Molecular and Cellular Endocrinology, 194(1-2), 147-155

Kuzmanov U, Kosanam H, Diamandis EP (2013) The sweet and sour of serological glycoprotein tumor biomarker quantification, Medicine, 11:31

Lange EM, Salinas CA, Zuhlke KA, Ray AM, Wang Y, Yunfei M, et al., (2012) Early onset prostate cancer has a significant genetic component, The Prostate, 72(2), 147-156

Lapthorn AJ, Harris DC, Littlejohn A, Lustbader JW, Canfield RE, Machin KJ, Morgan FJ, Isaacs NW (1994) Crystal structure of human chorionic gonadotropin, Nature. 9: 369(6480) 455-61

Levernæs MCS, Broughton MN, Reubsaet L, Halvorsen TG (2017) To elute or not to elute in immunocapture bottom-up LC-MS. J Chromatogr B Analyt Technol Biomed Life Sci.1055-1056:51-60

Li D, Zhang P, Li F, Chi L, Zhu D, Zhang Q, Chi L (2015) Recognition of N-Glycoforms in Human Chorionic Gonadotropin by Monoclonal Antibodies and Their Interaction Motifs, J Biol Chem. 11; 290(37): 22715–22723

Licht P, Fluhr H, Neuwinger J, Wallwiener D, Wildt L (2007) Is human chorionic gonadotrophin directly involved in the regulation of human implantation? Mollecular and Cellular Endocrinology, 15:85-92

Link-Lenczowski P, Bubka M, Balog CIA, Koeleman CAM, Butters TD, Wuhrer M, Lityńska A (2018) The glycomic effect of N-acetylglucosaminyltransferase III overexpression in metastatic melanoma cells. GnT-III modifies highly branched N-glycans. Glycoconj J. 35(2):217-231

Lis H and Sharon N. (1993) Protein glycosylation. Structural and functional aspects, Eur J Biochem. 15;218(1): 1-27

Liu X, Nie H, Zhang Y, Yao Y, Maitikabili A, et al. (2013) Cell Surface-Specific N-Glycan Profiling in Breast Cancer. PLOS ONE 8(8): e72704

Lochmüller CH, Reese C, Aschman AJ, Breiner SJ (1993) Current strategies for prediction of retention in high-performance liquid chromatography, Journal of Chromatography, 656 (1–2), 3-18

Lu J, Isaji T, Im S, Fukuda T, Kameyama A, Gu J (2016) Expression of N-Acetylglucosaminyltransferase III suppresses a2,3-sialylation, and its distinctive functions in cell migration are attributed to a2,6-sialylation levels. J. Biol. Chem. 291, 5708–5720

Lund H, Snilsberg AH, Paus E, Halvorsen TG, Hemmersbach P, Reubsaet L (2013) Sports drug testing using immuno-MS: clinical study comprising administration of human chorionic gonadotropin to males Anal. Bioanal. Chem. 405(5)

Lundin M, Nordling S, Lumdin J, Alfthan H, Stenman U, Haglund C (2001) Tissue expression of human chorionic gonadotropin beta predicts outcome in colorectal cancer: a comparison with serum expression. Int J Cancer. 95(1): 18-22

Lunenfeld B (2004) Historical perspectives in gonadotrophin therapy, Human Reproduction Update: 10 (6) 453–467

Lurain JR (2010) Gestational trophoblastic disease I: epidemiology, pathology, clinical presentation and diagnosis of gestational trophoblastic disease, and management of hydatidiform mole. AJOG 203 (6) 531-539

Lurain JR (2011) Gestational trophoblastic disease II: classification and management of gestational trophoblastic neoplasia: 204 (1), 11–18

Lustbader, JW, Birken, S., Pileggi, NF., Kolks, MAG (1989) Crystallization and characterization of human chorionic gonadotropin in chemically deglycosylated and enzymically desialylated states Biochemistry. 1989 Nov 28;28(24): 9239-43

Mangili G, Lorusso D, Brown J, Pfisterer J, Massuger L, Vaughan M, Ngan HY, Golfier F, Sekharan PK, Charry RC, Poveda A, Kim JW, Xiang Y, Berkowtiz R, Seckl MJ, (2014) Trophoblastic disease review for diagnosis and management: a joint report from the International Society for the Study of Trophoblastic Disease, European Organisation for the Treatment of Trophoblastic Disease, and the Gynecologic Cancer InterGroup, Int. J. Gynecol. Cancer 24 (9), S109–S116

Mass Spectrometry of Glycobiology Analysis Manual, 2nd Edition (Sigma Aldrich, 2016). https://www.sigmaaldrich.com/technicaldocuments/articles/biology/glycobiology/mass-spectrometry-of-glycans.html

Matsuura N, et al. (1998) Gene expression of fucosyl- and sialyl-transferases which synthesize sialyl Lewisx, the carbohydrate ligands for E-selectin, in human breast cancer. Int. J. Oncol. 12, 1157–1164

Matsuura S and Chen HC (1980) A simple and effective solvent system for elution of gonadotropins from concanavalin A affinity chromatography. Analytical biochemistry, 106(2), 402-410

Matsuura Y, Takehira M, Joti Y, Ogasahara K, Tanaka T, Ono N, Kunishima N, Yutani K, (2015) Thermodynamics of protein denaturation at temperatures over 100 °C: CutA1 mutant proteins substituted with hydrophobic and charged residues, Scientific Reports volume5, Article number: 15545

Maymó J, Pérez-Pérez A, Maskin B, Dueñas J, Calvo Juan C, Varone C (2012) The Alternative Epac/cAMP Pathway and the MAPK Pathway Mediate hCG Induction of Leptin in Placental Cells. 7(10)

McComb ME, Perlman DH, Huang H, Costello CE (2007) Evaluation of an on-target sample preparation system for matrix-assisted laser desorption/ionization time-of-flight mass spectrometry in conjunction with normal-flow peptide high-performance liquid chromatography for peptide mass fingerprint analyses, Rapid Commun. Mass Spectrom, 21: 44–58

Melchor A, Orlando Z, Rodríguez O, Zulueta Rodríguez O, Conde IB, Bequer Ariza D, et al. (2005) Purification of Human Chorionic Gonadotropin from Pregnancy Urine by Immunoaffinity Chromatography Using a Monoclonal Antibody Anti-β Chain hCG, Hybridoma, 24(5):258-62

Mello JB, Ramos Cirilo PD, Michelin OC, Custódio Domingues MA, Cunha Rudge MV, Rogatto SR, Maestá I. (2017) Genomic profile in gestational and non-gestational choriocarcinomas. Placenta. 50:8-15

Merz WE (1994) The primate placenta and human chorionic gonadotropin. Exp Clin Endocrinol. 102(3) 222-34

Mise T and Bahl OP (1981) Assignment of Disulfide Bonds in the β Subunit of Human Chorionic Gonadotropin, Journal of Biological Chemistry, 256:13, 6547-6592

Miwa HE, Song Y, Alvarez R, Cummings RD, Stanley P (2012) The bisecting GlcNAc in cell growth control and tumor progression. Glycoconj J. 29(8-9)

Mora J, Gascón N, Tabernero JM, Rodríguez-Espinosa J, González-Sastre F (1996) Different hCG assays to measure ectopic hCG secretion in bladder carcinoma patients Br J Cancer.74(7):1081-4

Moriwaki K and Miyoshi E (2010) Fucosylation and gastrointestinal cancer World J Hepatol. 27;2(4):151-61

Murdoch S, Djuric U, Mazhar B, Seoud M, Khan R, Kuick R, Bagga R, Kircheisen R, Ao A, Ratti B, Hanash S, Rouleau GA, Slim R (2006) Mutations in NALP7 cause recurrent hydatidiform moles and reproductive wastage in humans Nature Geneticsvolume 38, 300–302

Muthu M, Chun S, Gopal J, Anthonydhason V, Haga SW, Jacintha P, Devadoss A, Oh JW (2020) Insights into bioinformatic applications for glycosylation: Instigating an awakening towards applying glycoinformatic resources for cancer diagnosis and therapy. International Journal of Molecular Sciences, 21(24), 9336

Nagae M, Kanagawa M, Morita-Matsumoto K, Hanashima S, Kizuka Y, Taniguchi N., et al. (2016) Atomic visualization of a flipped-back conformation of bisected glycans bound to specific lectins. Sci. Rep. 6:22973

Nagai Y, Nakamoto T, Nakasone T, Taira, Y., & Aoki, Y (2018) High-risk gestational choriocarcinoma with an unusual presentation and the treatment course of refractory or quiescent/minimally invasive disease. Gynecologic oncology reports, 26, 56–59

Narasimhan S (1982) Control of glycoprotein synthesis. UDP-GlcNAc: glycopeptide beta 4-Nacetylglucosaminyltransferase III, an enzyme in hen oviduct which adds GlcNAc in beta 1–4 linkage to the beta-linked mannose of the trimannosyl core of N-glycosyl oligosaccharides. J Biol, Chem. 257(17):10235–10242

Nemansky M, Edzes HT, Wijnands RA, Van den Eijnden DH (1992) The polypeptide part of human chorionic gonadotrophin affects the kinetics of alpha 6-sialylation of its N-linked glycans but does not alter the branch specificity of CMP-NeuAc:Gal beta 1-4GlcNAc-R alpha 2-6-sialyltransferase; 2(2):109-17

Nemansky M, De Leeuw R, Wijnands RA, Van den Eijnden DH (1995) Enzymic remodelling of the N- and 0-linked carbohydrate chains of human chorionic gonadotropin. Effects on biological activity and receptor binding. Eur.J Biochem. 227, 880-888

Niimi K, Yamamoto E, Fujiwara S, Shinjo K, Kotani T, Umezu T, Kajiyama H, Shibata K, Ino K, Kikkawa F (2012) High expression of N-acetylglucosaminyltransferase IVa promotes invasion of choriocarcinoma, British Journal of Cancer 107, 1969–1977

Ning F, Hou H, Morse AN, Lash GE (2019) Understanding and management of gestational trophoblastic disease. F1000Research, 8, F1000 Faculty Rev-428

Nishino K, Yamamoto E, Niimi K, Sekiya Y, Yamashita Y, Kikkawa F (2017) N-acetylglucosaminyltransferase IVa promotes invasion of choriocarcinoma. Oncol Rep. 38(1):440-448

Nwabuobi C, Arlier S, Schatz F, Guzeloglu-Kayisli O, Lockwood CJ, Kayisli UA (2017) hCG: Biological Functions and Clinical Applications, Int. J. Mol. Sci. 18(10) 2037

Okuyama N, Ide Y, Nakano M, Nakagawa T, Yamanaka K, Moriwaki K, Murata K, Ohigashi H, Yokoyama S, Eguchi H, Ishikawa O, Ito T, Kato M, Kasahara A, Kawano S, Gu J, Taniguchi N, Miyoshi E (2006) Fucosylated haptoglobin is a novel marker for pancreatic cancer: a detailed analysis of the oligosaccharide structure and a possible mechanism for fucosylation. Int J Cancer, 118: 2803-2808

Oliveira-Ferrer L, Legler K, Milde-Langosch K. (2017) Role of protein glycosylation in cancer metastasis. Semin Cancer Biol. 44:141-152

O'Neil, M (2001) Merck Index : an encyclopedia of chemicals, drugs, & biologicals : 13th ed. United States: MERCK & CO INC

OpenStax (2013) Illustration from Anatomy & Physiology, Connexions Web site. http://cnx.org/content/col11496/1.6

Orczyk-Pawiłowicz M, Augustyniak D, Hirnle L, Kątnik-Prastowska I (2012) Degree of sialylation and fucosylation of plasma and amniotic immunoglobulin G changes progressively during normal pregnancy, Prenatal diagnosis, 32(5), 432-439

Ou L, He X, Liu N, Song Y, Li J, Gao L et al. (2020) Sialylation of FGFR1 by ST6GalI overexpression contributes to ovarian cancer cell migration and chemoresistance. Mol. Med Rep. 21, 1449–1460

Partridge EA, Le Roy C, Di Guglielmo GM, Pawling , Cheung P, Granovsky M, et al. (2004) Regulation of cytokine receptors by Golgi N-glycan processing and endocytosis. Science 306, 120–124

Patel R (2015) MALDI-TOF MS for the diagnosis of infectious diseases, Clin Chem. 61(1): 100-11

Peixoto A, Relvas-Santos M, Azevedo R, Santos LL, Ferreira J (2019) Protein Glycosylation and Tumor Microenvironment Alterations Driving Cancer Hallmarks,Front. Oncol., https://doi.org/10.3389/fonc.2019.00380

Pereira L (2008) Porous Graphitic Carbon as a Stationary Phase in HPLC: Theory and Applications, Journal of Liquid Chromatography & Related Technologies®, 31:11-12, 1687-1731

Pierce JG and Parsons TF (1981) Glycoprotein Hormones: Structure and Function, Annu. Rev. Biochem, 50(1), 465-495

Pinho SS and Reis CA (2015) Glycosylation in cancer: mechanisms and clinical implications, Nature Reviews, 15, 540-555

Pinho SS, Figueiredo J, Cabral J, Carvalho S et al., (2013) E-cadherin and adherens-junctions stability in gastric carcinoma: Functional implications of glycosyltransferases involving N-glycan branching biosynthesis, N-acetylglucosaminyltransferases III and V, Biochimica et Biophysica Acta (BBA) - General Subjects,1830 (3), 2690-2700

Raju, TS and Lang SE (2014) Diversity in structure and functions of antibody sialylation in the Fc. Curr. Opin. Biotechnol. 30, 147–152

Rao CV and Lei ZM (2007) The past, present and future of nongonadal LH/hCG actions in reproductive biology and medicine. Mol Cell Endocrinol. 15;269(1-2):2-8

Raposo CD, Canelas AB, Barros MT (2021) Human Lectins, Their Carbohydrate Affinities and Where to Find Them. Biomolecules, 11(2), 188

Reily C, Stewart TJ, Renfrow MB, et al. (2019) Glycosylation in health and disease. Nat Rev Nephrol 15, 346–366

Reis CA, Osorio H, Silva L, Gomes C, David L (2010) Alterations in glycosylation as biomarkers for cancer detection, J Clin Pathol. 63(4): 322-9

Rejlekova K, Cursano MC, De Giorgi U, Mego M (2019) Severe Complications in Testicular Germ Cell Tumors: The Choriocarcinoma Syndrome. Frontiers in endocrinology, 10, 218

Roberge S, Giguere Y, Villa P, et al. (2012) Early administration of low-dose aspirin for the prevention of severe and mild preeclampsia: a systematic review and meta-analysis. Am J Perinatol, 29:551–6

Rodrigues E and Macauley MS (2018) Hypersialylation in cancer: modulation of inflammation and therapeutic opportunities. Cancers 10, E207

Rodríguez AM, Rodríguez OZ, Conde IB, Del Pino YR, Pérez L, López Cisnero R, Duque M, et al. (2005) Purification of Human Chorionic Gonadotropin from Pregnancy Urine by Immunoaffinity Chromatography Using a Monoclonal Antibody Anti-β Chain hCG , Hybridoma 24(5)

Rombouts Y, Willemze A, van Beers JJBC, et al. (2016) Extensive glycosylation of ACPA-IgG variable domains modulates binding to citrullinated antigens in rheumatoid arthritisAnnals of the Rheumatic Diseases,75:578-585

Rosen A, Hermoso M, Cook-Andersen H, de Ziegler D (2013) Purification of Human Chorionic Gonadotropin (hCG) for Intrauterine Injection. PCRS Abstracts 99(3), Supplement

Ruhaak LR, Uh HW, Deelder AM, Dolhain RE, Wuhrer M (2014) Total Plasma N-Glycome Changes during Pregnancy. J. Proteome Res. 13, 1657–1668

Sasaki Y, Ladner DG, Cole LA (2018) Hyperglycosylated human chorionic gonadotropin and the source of pregnancy failures, Fertility and Sterility, 89:(6)

Schmidt SR (2017) Controlling Glycosylation in Fusion Protein Manufacturing to Generate Potent Biobetters, Biopress International, https://bioprocessintl.com/manufacturing/biosimilars/controlling-glycosylation-in-fusion-protein-manufacturing-to-generate-potent-biobetters/

Schneider M, Al-Shareffi E, Haltiwanger RS (2017) Biological functions of fucose in mammals. Glycobiology, 27:601–618

Scott DA and Drake RR (2019) Glycosylation and its implications in breast cancer. Expert review of proteomics, 16(8), 665–680

Seberger PJ, Chaney WG (1999) Control of metastasis by Asn-linked, β1-6 branched oligosaccharides in mouse mammary cancer cells. Glycobiology 9, 235–241

Seckl MJ (2009) Newlands ES. Chapter 12. Investigation and treatment of patients with persistent gestational trophoblastic disease and gestational trophoblastic tumours/neoplasia in the United Kingdom. In: Hancock BW, Seckl MJ, Berkowitz RS, Cole LA, eds. Gestational trophoblastic disease, 3rd edn.

Seckl MJ, Sebire NJ, Berkowitz RS. (2010) Gestational trophoblastic disease. Lancet. 28; 376(9742):717-29

Sethi MK, Thaysen-Andersen M, Smith JT, Baker MS, Packer NH, Hancock WS, et al. (2014) Comparative N-glycan profiling of colorectal cancer cell lines reveals unique bisecting GlcNAc and a-2, 3-linked sialic acid determinants are associated with membrane proteins of the more metastatic/aggressive cell lines. J Proteome Res. 13:277–88

She YM, Tam RY, Li X, Rosu-Myles M, Sauvé S (2020) Resolving Isomeric Structures of Native Glycans by Nanoflow Porous Graphitized Carbon Chromatography–Mass Spectrometry Analytical Chemistry 92 (20), 14038-14046

Shevchenko A, Tomas H, Havli J, et al. (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nat Protoc 1, 2856–2860

Shigeta M, Shibukawa Y, Ihara H, Miyoshi E, Taniguchi N, Gu J (2006) b1,4-N-Acetylglucosaminyltransferase III potentiates b1 integrinmediated neuritogenesis induced by serum deprivation in Neuro2a cells. Glycobiology 16, 564–571

Shimizu H, Ochiai K, Ikenaka K,Mikoshiba K, Hase S (1993) Structures of N-linked sugar chains expressed mainly in mouse brain. J. Biochem. 114:5

Singh R and Choi BK (2019) Siglec1-expressing subcapsular sinus macrophages provide soil for melanoma lymph node metastasis. eLife 8, e48916

Singh SK, Nage N, Jagani H, Maiti M, Ranbhor RS (2018) Glycan mapping of recombinant human follicle-stimulating hormone by mass spectrometry Rep Biol 18:380-384

Sivalingam VA, Duncan WC, Kirk E, Shephard LA, Horne AW (2012) Diagnosis and management of ectopic pregnancy, J Fam Plann Reprod Health Care. 37(4): 231–240

Śliwa A, Kubiczak M, Szczerba A, Walkowiak G, Nowak-Markwitz E, Burczyńska B, Butler S, Iles R, Białas P, Jankowska A (2019) Regulation of human chorionic gonadotropin beta subunit expression in ovarian cancer, BMC Cancer, 19:746

Smitz J, Platteau P (2020) Influence of human chorionic gonadotrophin during ovarian stimulation: an overview. Reprod Biol Endocrinol. 6;18(1):80

Song Y, Aglipay JA, Bernstein JD, Goswami S, Stanley P (2010) The bisecting GlcNAc on N-glycans inhibits growth factor signalling and retards mammary tumour progression. Cancer Res. 70:3361–3371

Stanley P and Cummings (2017) Structures common to different glycans. Essentials of Glycobiology [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015–2017. Chapter 14

Stanley P, Schachter H, Taniguchi N. N-Glycans (2009) In: Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME, editors. Essentials of Glycobiology. 2nd ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press Chapter 8

Stanley P, Taniguchi N, Aebi M (2017) N-Glycans. In: Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar RL, Seeberger PH, editors. Essentials of Glycobiology [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015–2017. Chapter 9

Stenman UH (2013) Determination of human chorionic gonadotropin, Best Practice & Research Clinical Endocrinology & Metabolism 27, 783–793

Stenman UH, Tiitinen A, Alfthan H, et al. (2006) The classification, functions and clinical use of different isoforms of hCG. Hum Reprod;12: 769–84

Stowell SR, Ju T, Cummings RD (2015) Protein glycosylation in cancer. Annual review of pathology, 10, 473–510.

Swindall AF and Bellis SL (2011) Sialylation of the fas death receptor by ST6Gal-I provides protection against fas-mediated apoptosis in colon carcinoma cells. J. Biol. Chem. 286, 22982–22990

Szájli E, Fehér T and Medzihradszky KF (2008). Investigating the quantitative nature of MALDI-TOF MS. Mol Cell Proteomics 7(12):2410-8

Szekeres-Barthó J (2002) Immunological relationship between the mother and fetus. Int. Rev. Immunol. 21, 471–495

Szymendera JJ (1986) Clinical usefulness of three monoclonal antibody-defined tumor markers: CA 19-9, CA 50, and CA125. Tumour Biol, 7: 333-342

Takegawa Y, Deguchi K, Keira T, Ito H, Nakagawa H, Nishimura S (2006) Separation of isomeric 2-aminopyridine derivatized N-glycans and N-glycopeptides of human serum immunoglobulin G by using a zwitterionic type of hydrophilic-interaction chromatography. Journal of Chromatography A, 1113, 177–181

Tan Z, Wang C, Li X, Guan F (2018) Bisecting N-acetylglucosamine structures inhibit hypoxia-induced epithelial-mesenchymal transition in breast cancer cells. Front. Physiol. 9:210

Taniguchi N and Korekane H (2011) Branched N-glycans and their implications for cell adhesion, signaling and clinical applications for cancer biomarkers and in therapeutics, BMB Reports. Korean Society for Biochemistry and Molecular Biology

Taylor M and Drickamer K (2011) Introduction to Glycobiology, Oxford University Press, 125-138

The American Collage of Obstetrician and Gynecologists (2002) Avoiding inappropriate clinical decisions based on false-positive human chorionic gonadotropin test results, ACOG No 278, 100: 1057-9

Ticconi C, Zicari A, Belmonte A, Realacci M, Rao CV, Piccione E (2007) Pregnancy-promoting actions of HCG in human myometrium and fetal membranes. Placenta 28,137-143

Toll H, Berger P, Hofmann A, Hildebrandt A, Oberacher H, Lenhof HP, Huber CG (2006) Glycosylation patterns of human chorionic gonadotropin revealed by liquid chromatography-mass spectrometry and bioinformatics, Electrophoresis, 27, 2734-2746

Trinchera M, Aronica A, Dall'Olio F (2017). Selectin Ligands Sialyl-Lewis a and Sialyl-Lewis x in Gastrointestinal Cancers. Biology, 6(1), 16

Trowsdale J and Betz AG (2006) Mother's little helpers: mechanisms of maternal-fetal tolerance. Nature Immunol. 7, 241–246

Tu CF, Wu MY, Lin YC, Kannagi R, Yang RB (2017) FUT8 promotes breast cancer cell invasiveness by remodelling TGF-β receptor core fucosylation. Breast Cancer Res, 19:1–15

Valmu L, Alfthan H, Hotakainen K, Birken S, Stenman (2006) Site-specific glycan analysis of human chorionic gonadotropin -subunit from malignancies and pregnancy by liquid chromatography--electrospray mass spectrometry. Glycobiology. 16(12): 1207-1218

van de Wall S, Santegoets KCM, van Houtum EJH, Büll C, Adema GJ (2020) Sialoglycans and siglecs can shape the tumor immune microenvironment. Trends Immunol. 41, 274–285

Varki A and Kornfeld S (2017) Historical Background and Overview. 1-18 In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, Chapter 1.

Varki A and Lowe JB (2009) Essentials of Glycobiology. 2nd edition., Chapter 6 Biological Roles of Glycans, The Consortium of Glycobiology Editors, La Jolla, California.

Vreeker GCM and Wuhrer M (2017) Reversed-phase separation methods for glycan analysis, Anal Bioanal Chem, 409: 359–378

Vuong PN, Guillet, J.L., Houissa-Vuong, S., Lhommé, C., Proust, A., Cristalli, B., (2000) Pathology of gestational trophoblastic tumors, Gynecol Obstet Fertil. 28(12): p. 913-926

Wang JR, Gao WN, Grimm R. et al. (2017) A method to identify trace sulfated IgG N-glycans as biomarkers for rheumatoid arthritis. Nat Commun 8, 631

Wang X, Gu J, Ihara H, Miyoshi E, Honke K, Taniguchi N (2006) Core fucosylation regulates epidermal growth factor receptor-mediated intracellular signaling. J Biol Chem, 281: 2572-2577

Wide L, Lee JY, Rasmussen C (1994) A change in the isoforms of human chorionic gonadotropin occurs around the 13th week of gestation. J. Clin. Endocrinol. Metab. , 78, 1419–1423

Woldemariam GA and Butch AW (2014) Immunoextraction-tandem mass spectrometry method for measuring intact human chorionic gonadotropin, free β-subunit, and β-subunit core fragment in urine. Clin Chem. 2014 Aug;60(8):1089-97

Xing D, Zheng G, Pallavajjala A, Schoolmeester KJ, LiuY, Haley L, Hu Y, Liu L, Logan L, Lin Y, Pearce KE, Sattler CA, TsaiY, Vang R, Hung C, Wu T, Ronnett BM (2019) Lineage-Specific Alterations in Gynecologic Neoplasms with Choriocarcinomatous Differentiation: Implications for Origin and Therapeutics Clin Cancer Res (25) (14) 4516-4529

Xing Y, Williams C, Campbell RK, Cook S, Knoppers M, Addona T, Altarocca V, Moyle WR (2001) Threading of a glycosylated protein loop through a protein hole: implications for combination of human chorionic gonadotropin subunits, Protein Sci, 10: 226-235

Xu B, Zhu Z, Tang Y, Huang Q, Wang X (1999) Purification of human chorionic gonadotrophin from urine by membrane filtration affinity chromatography with a positively charged membrane. Protein Expr Purif;16(2):221-3

Yang S and Zhang H (2012) Solid-phase glycan isolation for glycomics analysis. Proteomics Clin Appl., 6(0): 596–608

Yokotani KT, Taniguchi R, Nakagawa T, Isobe T, Yoshimura M, Tsubota N, Hasegawa K, Ohsawa N, Baba S, Yasui H, Nishimura R (1997) Expression of alpha and beta genes of human chorionic gonadotropin in lung cancer Int J Cancer, 16;71(4): 539-44

Young C, Condina MR, Briggs MT, Moh ES, et al. (2021) In-House Packed Porous Graphitic Carbon Columns for Liquid Chromatography-Mass Spectrometry Analysis of N-Glycans, Frontiers in Chemistry 9

Zhang L and Hagen K (2019) O-Linked glycosylation in Drosophila melanogaster, Current Opinion in Structural Biology, 56, 139-145

Zhang W, Wang H, Zhang L, Yao J, Yang P (2011) Large-scale assignment of N-glycosylation sites using complementary enzymatic deglycosylation, Talanta, 85(1), 499-505

Zhao Y, Itoh S, Wang X, Isaji T, Miyoshi E, Kariya Y, Miyazaki K, Kawasaki N, Taniguchi N, Gu J (2006) Deletion of core fucosylation on alpha3beta1 integrin down-regulates its functions. J Biol Chem, 281: 38343-38350

Zhao Y, Xu X, Fang M, Wang H, You Q, Yi C, Ji J, Gu X, Zhou P, Cheng C, Gao C (2014) Decreased Core-Fucosylation Contributes to Malignancy in Gastric Cancer PLoS ONE 9(4)

Zheng X, Price CA, Tremblay Y, Lussier JG, Carrière PD (2008) Role of transforming growth factor-beta1 in gene expression and activity of estradiol and progesterone-generating enzymes in FSH-stimulated bovine granulosa cells. Reproduction., 136(4):447-57